

PERMANENT COUNCIL FOR THE ORGANIZATION
OF INTERNATIONAL CONGRESSES
FOR PHONETIC SCIENCES

Established in 1932 at Amsterdam

President	Prof. D. B. Fry	London	England
Vice-Presidents	Prof. B. Hála	Praha	Czechoslovakia
	Prof. M. Halle	Cambridge (Mass.)	USA
General Secretary	Prof. R. Gsell	Grenoble	France
Honorary Members	Prof. M. Cohen	Viroflay	France
	Prof. M. P. Fouché †	Paris	France
	Prof. R. Jakobson	Cambridge (Mass.)	USA
	Prof. D. Jones †	Gerrards' Cross	England
	Prof. L. Kaiser	Amsterdam	Netherlands
	Prof. A. de Lacerda †	Coimbra	Portugal
	Prof. E. Zwirner	Köln	W. Germany
Members	Prof. V. A. Artemov	Moscow	USSR
	Prof. S. K. Chatterji	Calcutta	India
	Prof. P. Delattre	Santa Barbara (Calif.)	USA
	Prof. G. Fant	Stockholm	Sweden
	Prof. E. Fischer-Jørgensen	Kopenhagen	Denmark
	Prof. A. V. Isačenko	Praha	Czechoslovakia
	Prof. M. Kloster Jensen	Bergen	Norway
	Prof. B. Malmberg	Lund	Sweden
	Prof. A. Martinet	Paris	France
	Prof. P. Moore	Chicago	USA
	Prof. I. Ochiai	Nagoya	Japan
	Prof. M. Onishi	Tokyo	Japan
	Prof. W. Pée	Ghent	Belgium
	Prof. G. E. Peterson †	Ann Arbor (Mich.)	USA
	Prof. L. K. Pike	Ann Arbor (Mich.)	USA
	Prof. A. Rigault	Montréal	Canada
	Prof. A. Rosetti	Bucarest	Rumania
	Prof. A. Sovijärvi	Helsinki	Finland
	Prof. G. Straka	Strasbourg	France
	Prof. F. Trojan	Wien	Austria

ORGANIZATION COMMITTEE
ОРГАНИЗАЦИОННЫЙ КОМИТЕТ
COMITÉ D'ORGANISATION
ORGANISATIONSKOMITEE

President	Prof. Dr. B. Hála (Praha)	Phonetics
General Secretary	Prof. Dr. M. Romportl (Praha)	Phonetics
Vice-Presidents	Prof. Dr. E. Pauliny (Bratislava)	Slovak, Phonology
	Prof. Dr. A. V. Isačenko (Praha)	Slavonic Linguistics, Phonology
Secretaries	Dr. J. Ondráčková (Praha)	Phonetics
	Doc. Dr. P. Janota (Praha)	Phonetics
Treasurer	L. Bínová (Praha)	
Members	Prof. Dr. J. Bělič (Praha)	Czech and Slavonic Dialectology
	Doc. Dr. F. Daneš (Praha)	Czech, Linguistics
	Prof. Dr. J. Merhaut (Praha)	Acoustics
	Prof. Dr. K. Ohnesorg (Brno)	Phonetics, Pedolinguistics
	MUDr. R. Poch (Praha)	Radiology
	Prof. Dr. K. Sedláček (Praha)	Otolaryngology, Audiology
	Prof. Dr. M. Seeman (Praha)	Phoniatrics
	Prof. Dr. V. Skalička (Praha)	General and Ugro-Finnic Linguistics
	Doc. Dr. A. Skaličková (Praha)	Phonetics
Prof. Dr. M. Sovák (Praha)	Special Education and Logopedics	
Prof. Dr. J. Vachek (Praha)	English and Czech, Linguistics	

Secretariat Praha 1-Malá Strana, Valdštejská 14 (ČSSR)

CHAIRMEN — ПРЕЗИДЕНТЫ — PRÉSIDENTS —
PRESIDENTEN

CHAIRMEN OF THE PLENARY SESSIONS

7. 9. M. Cohen, E. Fischer-Jørgensen, B. Havránek
8. 9. G. Fant, A. Rigault, A. Sovijärvi
9. 9. P. Delattre, L. Kaiser, M. Onishi
10. 9. M. Kloster Jensen, B. Malmberg, W. Pée
11. 9. V. A. Artemov, A. Martinet, E. Zwirner

CHAIRMEN OF THE SECTION MEETINGS

- | | | | | | |
|--------|-------|--------|---|---|---|
| 7. 9. | 15.00 | Room A | O. S. Achmanova
J. Vachek
D F. Daneš
W. R. Lee | B L. A. Čistovič
H. Mol
E L. Handzel
K. Ohnesorg | C J. Bělič
K. Taranovski |
| | 16.50 | Room A | G. Hammarström
E. Hamp
D R. de Bray
R. Zinder | B R. Gsell
T. Tervoort
E F. Trojan | C G. Lotzmann
P. Trost |
| 8. 9. | 11.00 | Room A | J. D. Gendron
G. Ungeheuer
D R. Avanesov
S. Stojkov | B K. Sedláček
E H. Arndorfer
W. von Raffler-Engel | C H. Penzl
L. Zabrocki |
| 9. 9. | 11.00 | Room A | I. Blanchard
G. Nickel
D A. N. Nurmachanova
E. Pauliny | B H. Birnbaum
G. Lindner
E R. Becker
W. Merlingen | C H. Mangold
F A. Avram
I. Mahnken |
| 11. 9. | 11.00 | Room A | P. L. Garvin
H. Schultink
D A. Frinta
D. Worth | B M. Kingsbourne
E. Pulgram
E M. Sovák
B. Vallancien | C L. Lisker
J. Merhaut |
| 11. 9. | 15.00 | Room A | L. K. Pike
E. Sivertsen
D J. Hamm
A. Peco | B I. Lehiste
E J. Black
W. A. Hucklebery | C F. Kolmer
S. de Vriendt
F R. Mikus
M. Le Roy |
| | 16.50 | Room A | M. Boudreault
M. Schubiger | B E. Beyer
G. Heike | C A. Schneider
W. Tscheschner |

- | | | | | | |
|--------|-------|--------|---|---|--|
| | | | D R. Auty
W. Jassem | E A. M. Truby
A. Vanvik | F K. Albrow |
| 12. 9. | 11.00 | Room A | J. Goosens
H. Lüdtke
D R. Charbonneau
A. Valdman | B M. De Grève
E Y. Lebrun
T. Slama-Cazacu | C R. McDavid
F. Winkel
F S. Garmadi
P. S. Green |
| | 15.00 | Room A | K. Horálek
R. de Bray
D H. Fournier | B P. McCarthy
E S. Keramitčievski | C F. Falc'hun
M. Wajskop
F J. Carnochan
A. Hadj-Salah |
| | 16.50 | Room A | M. Romportl
K. S. Šaumjan
D O. Parlàngeli
M. Rossi | B A. C. Gimson
E M. Seeman
E. Vasiliu | C A. S. Abramson
G. F. Meier
F M. Spencer
J. Young |
| 13. 9. | 9.00 | Room A | B. van den Berg
S. Bergsveinsson
D A. Haudricourt
M. Iliescu | B Y. Shen
V. Skalička
E L. S. Harms | C R. Jacobson
G. Storms
F E. Carney
C. G. Söderberg |

Official opening of the Congress

(House of Artists, 7. 9. 1967)

INAUGURAL ADDRESS BY THE PRESIDENT OF THE CZECHOSLOVAK ACADEMY OF SCIENCES ACADEMICIAN FRANTIŠEK ŠORM

Ladies and Gentlemen,

It is a great privilege and honour for me to greet you cordially on behalf of the Czechoslovak Academy of Sciences and my own behalf at the VIth International Congress of Phonetic Sciences. The decision of the International Committee of Phonetic Sciences to hold the VIth Congress in Prague this year was welcomed with pleasure by the Czechoslovak scientific circles and I welcome here in you, who came to participate in the Congress, the outstanding representatives of phonetic sciences from all over the world.

A great number of international scientific meetings, important congresses and symposia has been held in Prague during the recent years under the auspices of the Czechoslovak Academy of Sciences. I should like to draw your attention to the fact that the XIIIth General Assembly of the International Astronomical Union ended here a short time ago and the XXIth Congress of the International Union of Pure and Applied Chemistry is being held here now. Your Congress is marked by one special feature. Scientists—specialists in a number of various scientific fields which, at the first glance, may appear to have very little in common, have gathered in Czechoslovakia to discuss the results of their research on spoken language and its perception research carried out from various points of departure with the application of different methods and sometimes directed towards various practical aims. And it is the very difference of approach which offers the possibility to compare and check the reliability of results in the particular field without the necessity to apply methods which are alien to such a field. It concerns a certain variation of the team work which is so widely advocated today, but on a large scale. It is the merit of the phonetic sciences that the inter-scientific exchange of experiences and co-operation has not started only today but that this path has been followed already since the First Congress held in Amsterdam 35 years ago.

The agenda of your Congress comprises lectures not only from the sphere of linguistics, phonetics and phonology, but, likewise, from a number of biological and medical sciences, otolaryngology, phoniatriy, physiology, roentgenology, psychology and pedagogical sciences—namely logopedics, technical sciences and among them primarily acoustics, not to mention those fields of sciences which present to the Congress

their results of application of phonetic sciences. Permit me, as the representative of the highest scientific institution in Czechoslovakia, which embodies all those branches of science, to express my satisfaction over the fact that the Czechoslovak science is likewise strongly represented in your deliberations. The fact that Prague and Czechoslovakia have been chosen to act as hosts to the Congress is, in my opinion, not a mere chance but I think that both tradition and current achievements of Czechoslovak phonetics and related scientific disciplines have also played their part in that decision. This concerns not only phonology which, as a part of structural linguistics, has developed and has been further theoretically enriched by Czechoslovak phoniatriy and logopedics, through the activities of the well-known Prague Linguistic Circle (Cercle linguistique de Prague), but it concerns also acoustics and scientific branches as well. If at present the Czechoslovak Academy of Sciences in its Commission of Acoustics and Sub-Commission of Language and Physiology Acoustics coordinates the work and applies the results of scientific research of all those branches of science it is certainly in accord with the trend stipulated by your Congresses.

The Czechoslovak Academy of Sciences and the Organizational Committee headed by prof. Hála, who was elected Chairman of the Prague Congress at the last Congress, have tried to create the best conditions for your sessions. Permit me, in conclusion, to wish you all great success in your deliberations and a pleasant sojourn in old Prague.

ALLOCUTION DU PRÉSIDENT DU CONGRÈS PROFESSEUR BOHUSLAV HÁLA

Très honorés Collègues, Mesdames et Messieurs!

En ma qualité de Président du 6^{ème} Congrès International des Sciences Phonétiques, élu au 5^{ème} Congrès à Münster, ainsi qu'au nom du Comité d'Organisation du 6^{ème} Congrès, je m'empresse de vous souhaiter à tous chaleureusement la bienvenue! C'est un grand plaisir pour nous que de pouvoir vous accueillir à Prague, vieille ville historique et capitale de notre République!

Nous vous avons invités à venir chez nous, et vous avez bien voulu répondre à cette invitation. Vous n'avez ni ménagé votre temps, ni reculé devant les inconvénients du voyage afin de pouvoir faire part, les uns aux autres, des résultats obtenus par vos recherches scientifiques accomplies au cours de ces dernières années dans vos Laboratoires et dans vos Instituts. Vous êtes arrivés des quatre coins du monde, de tous les continents de notre globe! En revanche, nous avons travaillé de notre mieux pour rendre votre séjour à Prague le plus agréable possible et nous espérons que vous en serez satisfaits.

Je suis convaincu que, d'un accord unanime, nous adresserons nos plus chaleureux remerciements à l'Académie Tchécoslovaque des Sciences pour avoir bien voulu patroner notre Congrès, et tout spécialement à son Président, Académicien François Šorm dont vous venez d'entendre l'allocution.

Nous remercierons également de leur présence M. le Professeur Ladislav Haňka, vice-ministre au Ministère de l'Education Nationale, ainsi que M. Antonín Otradovec, Maire-adjoint de la ville de Prague.

Merci enfin à tous ceux qui par leur présence à Prague et par leurs exposés rehausseront l'éclat et l'importance de notre Congrès.

Nous avons choisi comme sujet principal le *Signal acoustique du langage et sa perception*. Tout en attirant l'attention des savants en premier lieu sur l'aspect acoustique du langage, le sujet que nous avons choisi laissait pourtant une assez grande liberté à tous de le traiter des points de vue les plus variés comme le prouvent les titres des communications qui nous sont parvenus.

Le nombre des communications inscrites ayant largement dépassé les possibilités des heures mises à la disposition du Congrès, nous avons été obligés de réduire, d'une part, le nombre des conférences en séances plénières, d'autre part, à 10 minutes seulement la durée de chaque communication faite en section. Nous avons adopté ces

réductions dans le but d'éviter un trop grand nombre de sections et nous espérons que ces mesures, quelque draconniennes qu'elles puissent paraître, auront votre entière approbation.

Je vous souhaite à tous le plus grand succès dans vos travaux!

Chers Congressistes, Mesdames et Messieurs, vous qui venez de l'étranger, puisque vous vous trouvez en ce moment sur le sol de notre pays, je vous prierai de me permettre d'ouvrir le Congrès par quelques souvenirs historiques qui vous montreront que les intérêts phonétiques chez nous remontent fort loin dans le passé. Je n'abuserai de votre patience que pendant quelques minutes.

Je commencerai par le célèbre réformateur de l'Église Romaine, Jean Hus, brûlé vif comme hérétique à Constance en 1415. Il enseignait à l'Université de Prague nouvellement créée par l'empereur Charles IV dont elle porte le nom, en 1348 selon le modèle de l'Université de Paris. Jean Hus mérite, à bon droit, l'admiration de tout phonéticien rien que pour le petit traité qui, actuellement, porte le titre *De Orthographia Bohemica*, dans lequel son auteur se proposait d'adapter la graphie latine aux besoins de la prononciation du tchèque de son époque. Sans s'en rendre compte, il a propagé ainsi les principes d'une écriture phonétique fondée sur la graphie latine, mais adaptée à la prononciation tchèque au moyen de signes diacritiques. Ces signes, vous pouvez les découvrir aisément, sous forme de virgules et de crochets, sur n'importe quelle inscription tchèque lors de vos promenades dans les rues de Prague. Par sa réforme, Hus a rendu l'orthographe tchèque presque entièrement phonétique; de ses signes diacritiques, le plus important est celui qui marque, par une virgule au-dessus de la lettre, la longueur des voyelles, la distinction des voyelles brèves et longues représentant un des traits caractéristiques du phonétisme de notre langue.

Après Jean Hus, une autre grande personnalité de l'histoire de la phonétique tchèque, vient le célèbre membre d'une nouvelle secte religieuse, celle des Frères Tchèques, Jean Blahoslav. C'était un homme largement instruit, prêtre et musicien en même temps, esprit animé d'un profond intérêt pour les qualités phoniques du langage. Par opposition à Hus qui était surtout théoricien, Blahoslav s'intéresse aux problèmes pratiques de la prononciation des orateurs, de sorte qu'il est, dans une certaine mesure, le précurseur de cette branche de la phonétique que, de nos jours, certains de nos collègues allemands ont créée sous le nom de „Sprechwissenschaft“. Pendant des années, Blahoslav a assisté aux sermons de ses confrères-prédicateurs pour noter consciencieusement tous leurs vices de prononciation ainsi que leurs mauvaises habitudes contractées dans l'exercice de leur profession d'orateurs ecclésiastiques.

Par ce procédé, il a accumulé, en écoutant et en observant, un riche matériel, d'après lequel il a composé, en 1571, un ouvrage intitulé *Vices des prédicateurs*. "La prononciation de ces derniers, souligne-t-il, est dépourvue de tout soin et de tout entraînement. Ils ne possèdent aucune connaissance de ce qu'il faut faire pour bien articuler et ils n'ont même pas envie de l'apprendre. Leurs gestes sont assez souvent exagérés et, comme le disait déjà Erasme de Rotterdam, célèbre érudit et littérateur

du 16^e siècle: il y en a qui seraient prêts à faire bouger même les oreilles, si la nature ne les avait pas formées immobiles."

En comparant plusieurs langues qu'il connaissait, Blahoslav considère que la langue tchèque sonne de beaucoup plus agréablement que certaines autres. Elle serait facile à prononcer, car "elle ne nécessite ni tordre la bouche comme le font les Français, ni de produire des sons sifflés ou chuintés comme les Italiens, ni enfin de parler par le nez comme le font les Polonais qui, par un nasillement désagréable, altèrent leur prononciation."

Après Blahoslav, presque trois siècles plus tard, au 19^e, la phonétique passe des mains des ecclésiastiques à celles des physiologues. Les recherches phonétiques acquièrent alors un caractère plus scientifique et gagnent, en même temps, en profondeur. C'est le mérite de deux physiologues tchèques, Purkyně et Čermák, qui ont consacré bien des années aux études de ce genre.

Le premier d'entre eux, Jean Évangéliste Purkyně, est un savant connu à l'étranger (en France notamment sous le nom francisé de Purkinje). Il était professeur de physiologie d'abord en Allemagne et, depuis 1849, à Prague. Il a publié de 1835 à 1836 une importante étude intitulée (je traduis en français) *Recherches concernant la physiologie du langage humain*. Dans cette étude, il souligne la nécessité de connaître à fond la configuration et le fonctionnement des organes phonateurs de l'homme. Mais, il ne s'occupe pas uniquement de l'aspect génétique du langage, il prête aussi son attention à l'aspect acoustique. Pour la description des sons du langage, il met à profit les résultats obtenus par ses précurseurs, s'inspirant notamment du célèbre ouvrage de Johann Wolfgang von Kempelen, *Mechanismus der menschlichen Sprache nebst der Beschreibung einer sprechenden Maschine*, paru en 1791. En passant, il sera peut-être intéressant de noter que, Kempelen, originaire de Pressbourg (actuellement Bratislava, capitale de la Slovaquie), serait en quelque sorte notre compatriote.

Cependant Purkyně ne se contente pas de données d'autres chercheurs, il effectue des expérimentations personnelles quoi que d'une façon très primitive; ainsi, rien que par le toucher, il découvre que la cavité pharyngienne s'élargit considérablement pour les voyelles antérieures, phénomène confirmé depuis par les radiographies des voyelles tchèques, et généralement connu sous le nom d'*espace de Purkyně* — ou bien de Purkinje, si vous voulez.

Tout en parlant des faits phonétiques, Purkyně se sert du terme „physiologie“. Cet usage ne doit pas surprendre, car le terme de „physiologie“ était en vigueur à l'époque et même ultérieurement, comme le prouve le titre *Grundzüge der Physiologie* employé par l'éminent phonéticien allemand Eduard Sievers pour la 1^{ère} édition, en 1876, de son manuel de phonétique. Et ce n'est que la deuxième édition qui paraît sous le titre de *Grundzüge der Phonetik*.

L'autre physiologue tchèque, dont le nom vient d'être cité, Jean Népomucène Čermák (remarquez bien que c'est déjà le 4^e Jean de suite), s'occupait notamment de l'exploration des articulations produites par l'action des cordes vocales. Il se servait pour cela du laryngoscope dont l'invention est, d'habitude, attribuée au professeur

de chant espagnol, Manual García, mais dû probablement au chirurgien anglais Liston, auteur de l'ouvrage *Practical Surgery*, publié en 1840.

Le grand service rendu par Čermák à la méthode laryngoscopique, a été d'introduire l'emploi de la *lumière artificielle*. Ce nouveau procédé permettant de travailler indépendamment de la lumière solaire a beaucoup facilité ses recherches. Čermák a pu décrire minutieusement et faire représenter par des croquis très précis les diverses positions des cordes vocales qu'elles occupent pendant la formation des sons et des bruits produits dans le larynx. La phonétique tchèque a été enrichie notamment par sa description de la consonne *h* sonore pour laquelle il se produit simultanément et un bruit consonantique, et une sonorité rappelant celle des consonnes sonores. La position des cordes vocales pour cette consonne a causé, à Čermák, une vive surprise: pour lui, il s'agissait là d'un phénomène acoustique remarquable (ein merkwürdiges akustisches Phänomen, écrit-il) à la découverte duquel il ne s'attendait pas, mais dont le bien-fondé a été confirmé, en 1930, par des prises cinématographiques.

Vers la fin du 19^e siècle, une époque nouvelle se préparait par l'introduction, en phonétique, des méthodes instrumentales et de l'expérimentation. Cette époque a été inaugurée par les travaux de l'abbé Rousselot. Pour répondre aux besoins pressants de recherches phonétiques plus approfondies, ce savant français a créé, comme on sait, une discipline nouvelle à laquelle il a donné le nom de „phonétique expérimentale“, nom combattu dès le début par le phonéticien danois, Otto Jespersen.

Les nouveaux courants scientifiques, suscités par l'abbé Rousselot, ont bientôt trouvé leurs répercussions à Prague où ils ont été rapportés par le professeur Joseph Chlumský, ancien assistant de l'Abbé. Ayant étroitement collaboré avec Rousselot pendant quatre ans, il a utilisé les idées et les méthodes de son maître pour créer à Prague, en 1920, un Laboratoire de phonétique expérimentale, et cela sur le modèle de celui de Paris. De ce fait, la phonétique rousselotienne a trouvé sa continuation directe sur le sol de Prague.

Parmi toutes les méthodes, la préférence de Chlumský allait aux enregistrements kymographiques et à leur dépouillement. Les résultats de ses recherches dans ce domaine ont été publiés, en 1928, dans un volumineux ouvrage intitulé *La quantité, la mélodie et l'accent d'intensité en tchèque*, paru, en abrégé, en version française dans le 6^e tome de la Revue de Phonétique.

Cela ne veut pas dire, bien entendu, que d'autres méthodes aient été laissées de côté. Au Laboratoire de Chlumský, incorporé à la Faculté des Lettres de L'Université Charles à Prague, on a effectué des travaux de radiographie, de cinématographie, d'oscillographie, d'analyses acoustiques etc. Cette première étape se termine, en 1939, par le décès du professeur Chlumský, date qui coïncide néfastement avec le déclenchement de la 2^e guerre mondiale.

Et ce n'est qu'en 1945 que, après une pause forcée de six ans, j'ai pu reprendre les travaux de mon ancien chef et maître, Joseph Chlumský, pour transformer son Laboratoire de phonétique expérimentale en Institut de Phonétique et en faire le noyau des recherches et de l'enseignement principal phonétique.

Mesdames et Messieurs, j'arrive au terme de mon bref exposé dont le but unique a été de vous familiariser avec quelques points saillants de l'histoire de notre phonétique à partir du début du 15^e siècle. Pour terminer, je déclare, en ma qualité de Président, le 6^e Congrès International des Sciences Phonétiques ouvert.

INAUGURAL ADDRESS BY THE PRESIDENT
OF THE PERMANENT COUNCIL
PROFESSOR D. B. FRY

Mr. President, Ladies and Gentlemen,

it is my privilege to speak to you on behalf of the Permanent International Council for Phonetic Sciences. My first duty, and one which I undertake with the very greatest pleasure, is to welcome you all to this, the Sixth International Congress of Phonetic Sciences. The Permanent Council does indeed offer to all the Delegates and Members of the Congress a very warm welcome.

I am sure that you will allow me also to speak on your behalf and to say how sensible we all are of the deep debt of gratitude we already owe to our friends in Czechoslovakia who have so kindly invited us to hold the Congress here in Prague. The organisation of such an international congress entails an immense expenditure of time and effort, even to reach the stage we are at this morning of assembling the members for the opening meeting, to say nothing of the work involved in carrying through the programme. We would like Professor Hála and his colleagues to know that we are sincerely grateful to them for undertaking this formidable task.

There are many reasons why it is a particular pleasure to be able to welcome you to our congress in Prague. The first of them lies quite clearly in the city of Prague itself, a city famed for its beauty not only throughout Europe but in every part of the world; so much so that the beauty of Prague has become a legend to many people who have never had the good fortune to visit it. For those of us who may be seeing it for the first time, it is indeed somewhat overwhelming to find that the experience of the reality transcends even the legend.

But the fame of Prague rests not only upon its beauty; it has been for centuries a centre of scholarship, whose reputation has through the years remained unsurpassed among the seats of learning in Europe. By its long unbroken tradition of scholarship Prague has conferred the greatest benefits upon our civilisation and we feel ourselves honoured to be invited to meet in this historic place.

Yet we have even more specific reasons for our pleasure in meeting here than the academic tradition in a general sense. No one whose work lies in the field of the phonetic sciences could fail to experience what our French friends might call "a certain emotion" upon visiting the home of the "Prague School". It is now considerably more than thirty years since many of us first became aware of the studies being carried out by the brilliant and devoted band of scholars who formed that School. The research

which they carried out and the ideas which they developed permanently influenced the course of work in phonetics and phonology and today we recognize with deep gratitude the debt which we owe to them.

We do not, however, live in the past, no matter how strong its influence may be, and indeed we cannot. So we come to the essential reason why we are meeting here today and why it gives us all such particular pleasure, and that is that the great traditions of scholarship that I have spoken of are carried on so worthily by Professor Hála, Professor Romportl and by their many colleagues, too numerous to mention on such an occasion as this. It is by their kind invitation that we meet here in Prague; it is their devoted work that has made it possible and it is for their sake that we count it a particular privilege to hold the Sixth International Congress of Phonetic Sciences in the Charles University. We thank them most warmly for all their efforts and we know that through their organisation and under their guidance the congress will be a signal success.

ALLOCUTION DU PRÉSIDENT DE LA SOCIÉTÉ INTERNATIONALE DES SCIENCES PHONÉTIQUES PROFESSEUR BERTIL MALMBERG

Monsieur le Président, Mesdames, Mesdemoiselles, Messieurs !

L'organisation appelée depuis sa fondation en 1936 la Société internationale des sciences phonétiques — en anglais International Society of Phonetic Sciences, en allemand Internationale Gesellschaft für phonetische Wissenschaften — et dont l'organisation et les activités futures seront définies dans les détails au cours de ce congrès, a pour but de stimuler et de faciliter le travail scientifique et pédagogique dans le domaine de nos sciences dans tous leurs aspects et de fournir, dans la mesure du possible, des matériaux de documentation et d'information à ses membres et à ses associations affiliées, pour les tenir au courant et les mettre en rapport avec des collègues, des centres de recherche, etc. Cette tâche devient de plus en plus importante avec l'évolution, dans les directions les plus diverses — parfois même inattendues — des sciences qui s'occupent de l'expression du langage humain sous sa forme sonore. Notre organisation, qui a repris ses activités sous une forme plus développée au congrès de 1961, tient à cette occasion à exprimer au comité d'organisation du sixième congrès international, inauguré ce matin à Prague, sa profonde reconnaissance pour avoir pris à sa charge la lourde tâche d'arranger un congrès mondial de dimensions qui dépassent de beaucoup celles des congrès précédents.

Le choix de Prague comme lieu du congrès était significatif. Le nom de Prague est intimement lié à un mouvement scientifique qui a joué un rôle plus grand, j'ose le dire et je suis prêt à défendre mon opinion, dans l'histoire des sciences phonétiques qu'aucun autre. L'école de Prague et les idées dites pragoises impliquent la création d'une base linguistique plus solide, plus réaliste, des recherches phonétiques, une définition de la phonétique dans le cadre de la linguistique générale, un point de départ des mesures physiologiques et acoustiques qui avait manqué aux efforts instrumentaux du début du siècle. Cette tradition est restée vivante à Prague. Inutile de mentionner les grands noms que tout le monde connaît.

Mais cette tradition linguistique s'est vue combinée dès le début et de façon heureuse à une autre, également importante: la recherche sérieuse et compétente sur la substance phonique sans laquelle la transmission linguistique ne fonctionnerait pas — la tradition d'un Chlumský, continuée par notre président M. Hála et par tous les collaborateurs et élèves qui lui doivent leur formation scientifique. Nous voyons dans la personne de M. Hála le 'altmeister' de cette tradition de phonétique ins-

trumentale, et nous rendons en ce moment hommage à ce qu'il représente, par ses travaux personnels, par les centres qu'il a dirigés pendant si longtemps, et par l'influence qu'il a eue.

Nous rendons hommage à tous ces savants, représentants de théories linguistiques et de recherches instrumentales en exprimant à notre président et à ses collègues et collaborateurs du comité toute notre reconnaissance pour nous avoir donné cette occasion de nous réunir dans la belle ville de Prague pour contribuer, par nos discussions et nos rapports, aux progrès des sciences phonétiques — sciences du langage humain.

МЕТОД ИССЛЕДОВАНИЯ РЕШАЮЩИХ ПРАВИЛ, ПРИМЕНЯЕМЫХ ПРИ ВОСПРИЯТИИ РЕЧИ

Л. А. ЧИСТОВИЧ*

Задача исследования распознавания фонем человеком состоит, в основном, в установлении акустических параметров речевого сигнала, измеряемых слухом, и в выявлении решающих правил, на основании которых звук определяется как та или иная фонема.

Первым шагом к решению этой задачи является определение и строгое количественное описание фонемных категорий. Под фонемной категорией понимается множество стимулов, воспринимаемых как одна и та же фонема. Акустический сигнал удобно рассматривать как точку в многомерном пространстве акустических параметров. Тогда фонемная категория будет соответствовать области в этом многомерном пространстве.

Очевидно, что для того, чтобы описать фонемную область, нужно указать, где находятся ее границы. Если граница между некоторыми двумя фонемами есть функция нескольких акустических параметров, она может иметь вид гиперповерхности. Логично предположить, что одна такая гиперповерхность будет разделять не только две частные фонемы, но две группы фонем, например, глухие и звонкие или мягкие и твердые. Тогда эта гиперповерхность будет отражать решающее правило, позволяющее перейти от акустических параметров к дифференциальному признаку: если сигнал находится по одну сторону гиперповерхности, ему присваивается одно значение по дифференциальному признаку, если он находится по другую сторону — другое значение. Мне кажется, что такое понимание дифференциального признака как решающей границы, разделяющей два класса фонемных категорий в пространстве акустических параметров — весьма близко к тому, которое дал Г. Фант в своей недавней статье „Природа дифференциальных признаков“ (1).

Итак, задача состоит в том, чтобы экспериментально определить положение границ между фонемными категориями в пространстве акустических параметров.

В принципе эту задачу можно решать обычным методом исследования фонемной идентификации синтетических речевых стимулов. Так она

* Институт физиологии им. И. П. Павлова АН СССР, Ленинград.

обычно и решается, если определяется граница между двумя фонемами по одному акустическому параметру. Однако, уже для трехмерного случая оказывается необходимым заранее записать на пленку и предъявлять для прослушивания так много различных стимулов, что работа становится очень трудно выполнимой. Дополнительная сложность состоит в том, что результаты идентификации стимулов разными испытуемыми нельзя объединять, т. к. положение границ у разных людей может не совпадать. Отсюда следует, что необходимо найти какой-то более быстрый и экономный метод определения границ между фонемными категориями.

В моем докладе я хочу рассказать об одном таком более быстром методе и привести, в качестве иллюстрации, некоторые результаты, полученные с его помощью.

Этот метод может быть назван методом активного поиска фонемной границы. Он является вариантом широко известного в психоакустике метода установки. Суть его состоит в том, что испытуемый сам управляет синтезатором сигналов, изменяя характер стимула, который он слушает. Испытуемому дается задание вращать только одну ручку прибора. Повернув ручку, испытуемый должен затем включить сигнал, прослушать его, снова повернуть ручку и продолжать это до тех пор, пока не удастся найти точку перехода от одной фонемы к другой. В простейшем случае ручка управления может быть связана только с одним параметром стимула. Остальные параметры стимула устанавливаются экспериментатором и остаются фиксированными во время поиска границы. Можно также одной ручкой управления изменять одновременно несколько параметров стимула, для чего экспериментатор должен задать между этими параметрами определенную функциональную зависимость. Поворот ручки будет соответствовать перемещению точки, представляющей сигнал, по некоторой траектории в пространстве акустических параметров.

На рис. 1 влево приведены для примера две траектории в пространстве двух параметров. Одна из них соответствует случаю, когда ручка связана только с одним параметром стимула, значение стимула по второму пара-

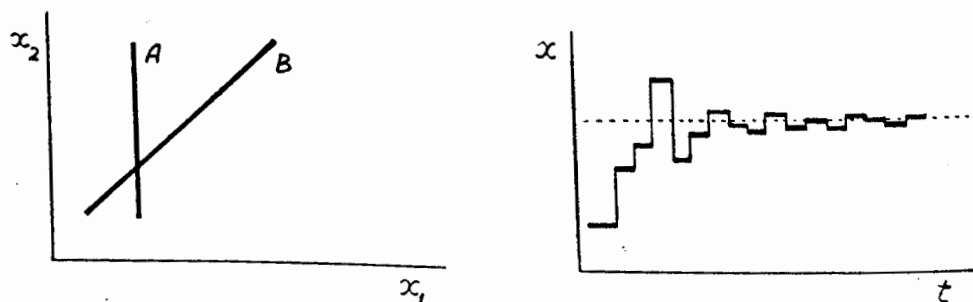


Рис. 1. Иллюстрации к методу активного поиска фонемной границы. Объяснения в тексте.

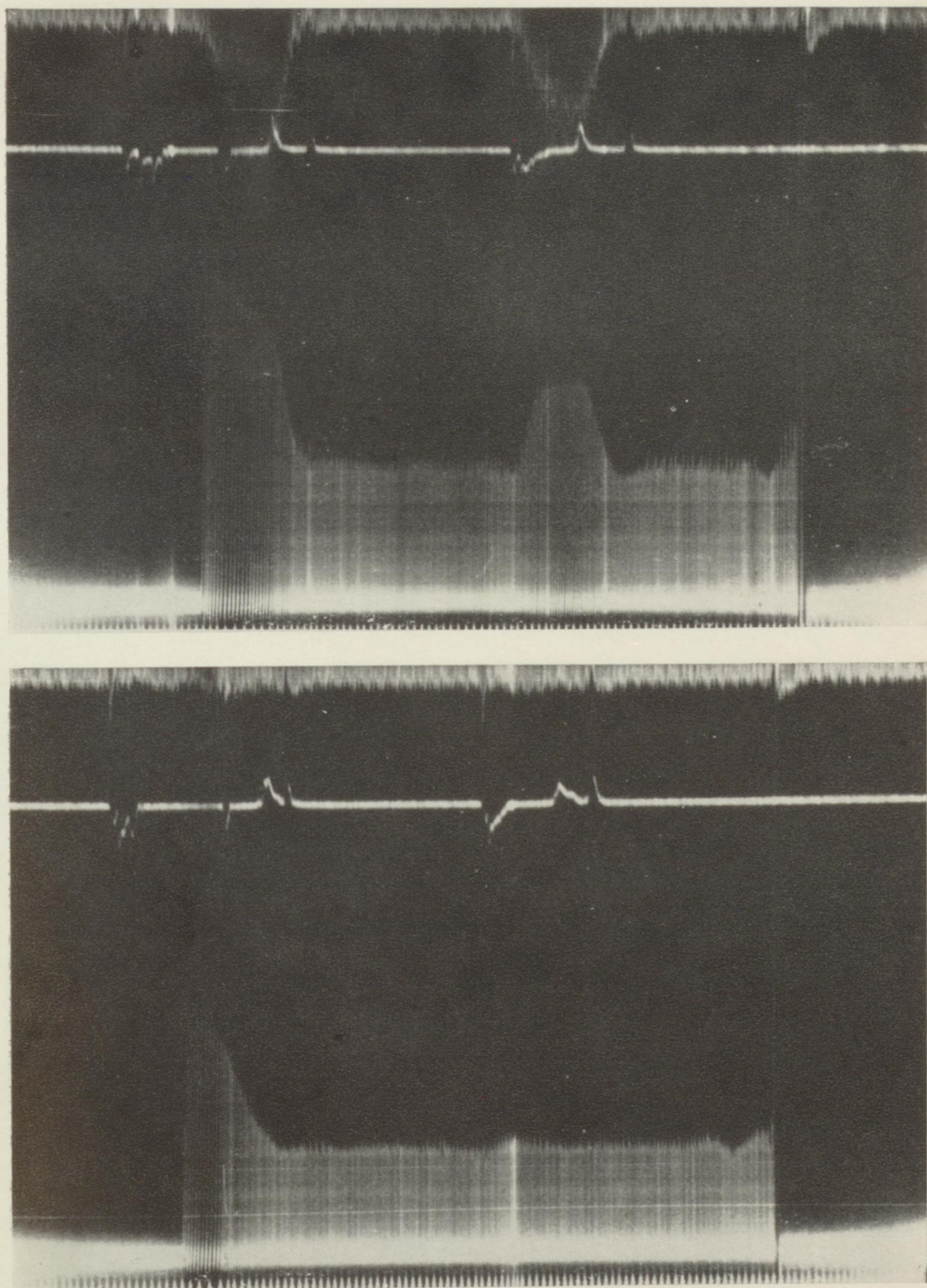


Рис. 2. Изменения периода основного тона (А) и внутриротового давления (Б) на протяжении слов *баба* и *мама*.

метру фиксировано. Другая траектория соответствует случаю, когда одновременно управляются оба параметра, связанные в данном примере линейной зависимостью.

На рис. 1 вправо показано, как во времени происходит процесс поиска границы. Его особенностью является то, что пока стимул находится в пределах одной фонемной категории, испытуемый изменяет значения параметра крупными шагами. Как только стимул пересекает фонемную границу, шаг изменения параметра резко сокращается. Испытуемый многократно проходит границу в ту и другую сторону, стремясь определить ее положение как можно точнее. Опыт показывает, что для одного определения границы испытуемому требуется прослушать 30—50 стимулов, на что уходит не более одной — двух минут.

После того, как испытуемый нашел границу и прекратил поиск, экспериментатор записывает значения параметров стимула. Для этого необходимо, чтобы в синтезаторе имелись достаточно точные индикаторы, с которых можно было бы прямо считывать значения параметров стимула.

Определение границы следует повторять несколько раз. В качестве окончательного результата эксперимента можно использовать среднее значение положения границы и среднее квадратическое отклонение отдельных определений. Последний показатель также представляет значительный интерес, т. к. он несомненно отражает точность слухового измерения параметра и стабильность критерия, применяемого при принятии решения.

Дальше я приведу некоторые данные, полученные методом поиска границ. Первая попытка применить этот метод была сделана в лаборатории Фанта — с помощью синтезатора OVE-1b мы пытались определить границы между гласными в двухформантной плоскости (2). Дальнейшие эксперименты были выполнены в Ленинграде с помощью установки, позволяющей получать очень маленький набор речеподобных стимулов. Синтезатор СГ слогов с ручным управлением находится пока в стадии изготовления. Основной целью экспериментов было накопление опыта по работе с данным методом.

Исследовалась граница между фонемами *б* и *м* в синтетических слогах *ба* и *ма*. Слог образовывался как последовательность из двух непосредственно примыкающих друг к другу „стационарных“ звуков. Источниками сигналов служили два генератора сложных гармонических напряжений, подключенные на вход двухканального электронного ключа. Преимуществом такого синтеза, при всей его примитивности, является возможность точного задания основной частоты, спектра, длительности и суммарной интенсивности для каждого из двух сигналов в последовательности.

Известно, что естественные *ба* и *ма* отличаются друг от друга по ряду акустических признаков. Одним из таких признаков, который, как ранее предполагалось, не используется человеком при восприятии, является

динамика изменения основного тона. На рис. 2 показано изменение периода основного тона из протяжении монотонно произнесенных слов *баба* и *мама*. Можно видеть, что на интервокальном *б* период основного тона увеличивается вместе с повышением внутриротового давления (верхняя кривая на нижнем снимке); при переходе к гласному он скачком возвращается к прежнему уровню. На интервокальном *м* частота основного тона или не меняется или изменяется очень незначительно.

В опытах с синтетическими слогами было сначала проверено предположение, что скачек основного тона при переходе от согласного к гласному используется для различения *б* и *м*. Исходные стимулы подбирались так, чтобы при одинаковой основной частоте на участках гласного и согласного они максимально напоминали слог *ма*. Затем стимулы изменялись так, что или понижалась основная частота на согласном или повышалась основная

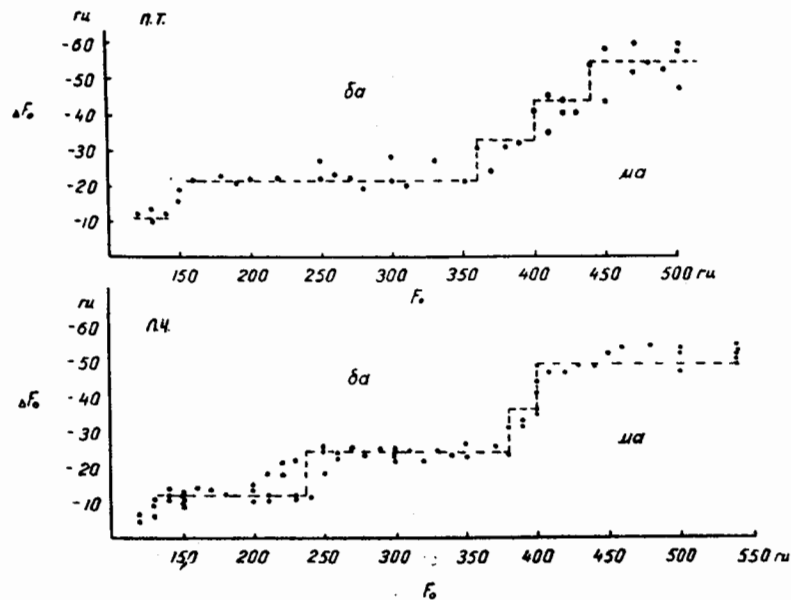


Рис. 3. Граница между *ба* и *ма* по признаку изменения основной частоты. Данные 2х испытуемых.

частота на гласном. Оказалось, что при этом стимулы начинали восприниматься как *ба*. Далее подробно исследовалось на четырех испытуемых какое значение изменения частоты при переходе от согласного к гласному соответствует границе между *м* и *б*. Измерения производились в широком диапазоне основных частот (3). Данные двух испытуемых приведены на рис. 3. По оси абсцисс отложены значения основной частоты на участке гласного в синтетическом слоге. По оси ординат — величина прироста основной частоты при переходе от согласного к гласному. В этих опытах

основная частота на гласном задавалась экспериментатором. Испытуемый изменял значение основной частоты на участке согласного, ища границу между *м* и *б*. Каждая точка на графике представляет среднее арифметическое из 5—10 определений границы. Видно, что величина изменения основной частоты, соответствующая границе, растет с увеличением абсолютного значения основной частоты, однако зависимость не является монотонной, а скорее имеет ступенчатый характер. Среднее квадратическое отклонение отдельных определений границы весьма мало на интервале, соответствующем середине ступеньки (3—4 гц), на краях оно резко возрастает.

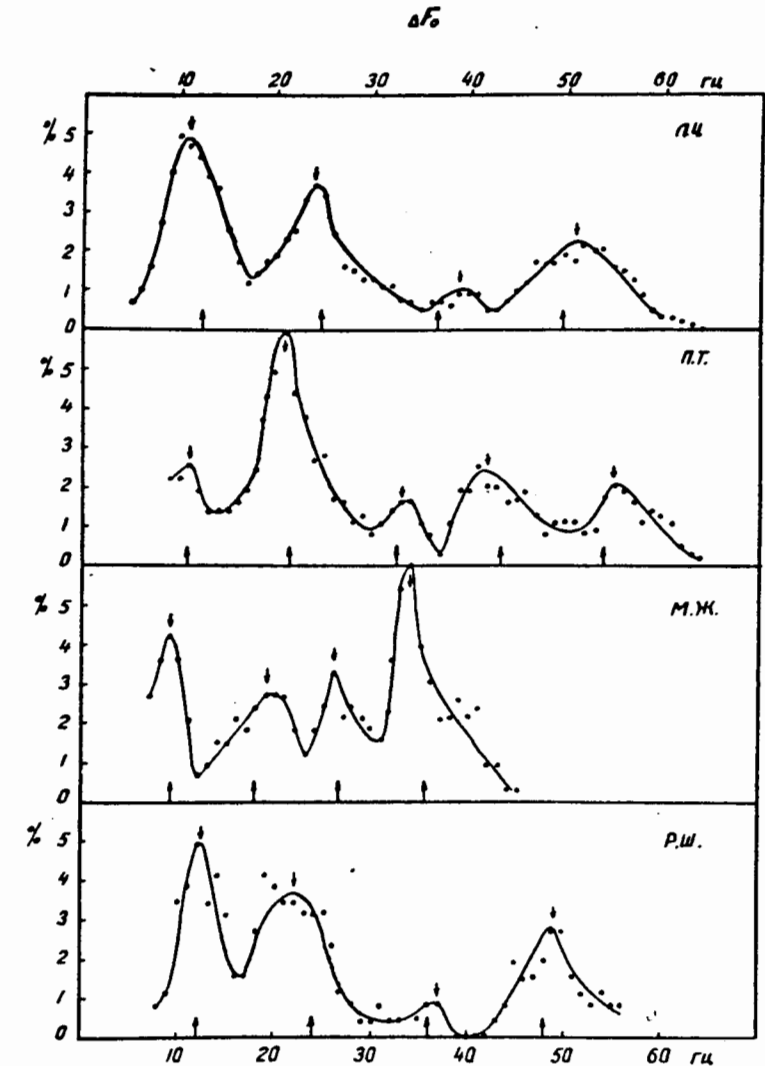


Рис. 4. Суммарные распределения определений границы между *ба* и *ма*. Данные 4х испытуемых.

Интересно, что значение ΔF_0 на второй ступеньке оказалось вдвое большим значения ΔF_0 на первой ступеньке. Это заставило обратиться к суммарным распределениям всех отдельных определений границы, полученных в опытах. Они приведены на рис. 4. Видно, что распределения имеют несколько пиков, расположенных примерно на равных расстояниях друг от друга. Причем это расстояние совпадает со значением ΔF_0 для первого пика.

Полученные данные можно объяснить с помощью представления о временном механизме измерения основной частоты звука слуховой системой. Предположим, что измеряемой мозгом величиной является интервал времени между соседними нервными импульсами, который, в зависимости от частоты сигнала равен $1, 2 \dots, k$ периодам основного тона. (Известно, что благодаря рефрактерности первый слуховой нейрон работает как делитель частоты.) Тогда для всего диапазона основных частот достаточно допустить одно значение порога, разделяющего *м* и *б*, однако это значение относится уже к частоте импульсации, а не прямо к частоте сигнала.

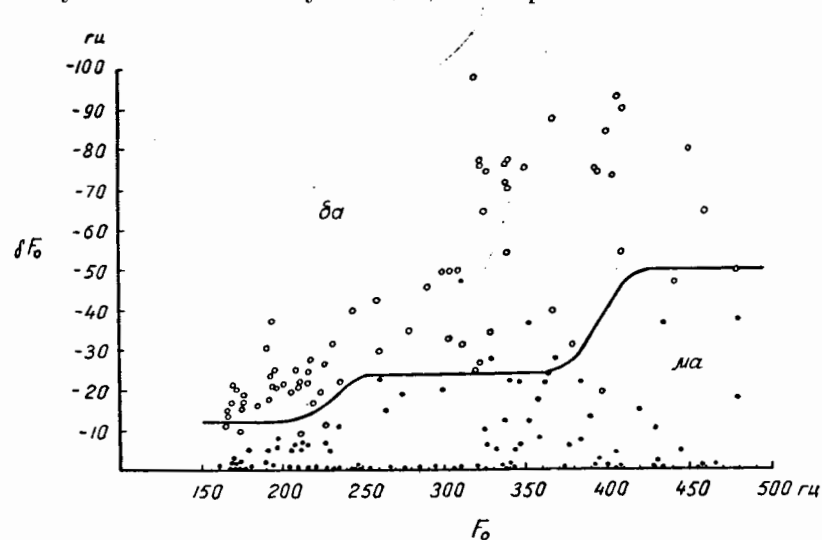


Рис. 5. Распределение естественных *ба* и *ма* в пространстве.

Очевидно, что исследование восприятия и исследование характеристик естественных звуков речи должны дополнять друг друга. После того, как в опытах с синтезом найдена граница между фонемами, которую использует человек при восприятии, целесообразно проверить, насколько хорошо эта граница разделяет естественные звуки речи.

Для проверки эффективности границ по изменению основной частоты, был использован следующий материал. Шесть дикторов наговорили монотонно на разных основных частотах слова *мама* и *баба*. Измерялись основная частота на втором гласном и изменение основной частоты при переходе

от интервокального согласного к гласному. На рис. 5 точками показаны данные касающиеся *м*, кружками — *б*. Кривая соответствует границе, определенной у одного из испытуемых. Можно видеть, что эта граница достаточно хорошо разделяет естественные *м* и *б* (правильное разделение в 94 % случаев). Худший из четырех полученных вариантов границы обеспечивал правильное разделение только в 83 % случаев.

Отсюда следует довольно тривиальный вывод, что изменение основной частоты не может быть единственным полезным признаком, различающим *б* и *м*.

Дальше я хочу рассказать о попытке использовать метод поиска границ для выяснения вопроса о том, как при восприятии объединяется информация, заключенная в различных акустических признаках сигнала, но имеющая один и тот же смысл.

Известно, что любые две фонемы различаются друг от друга более, чем по одному акустическому признаку. Казалось бы рациональным, чтобы граница между фонемами была в этом случае функцией от всех этих полезных акустических параметров. Иначе говоря, граница между фонемами А и В по параметру x_1 , должна, казалось бы, зависеть от того, какие значения имеет сигнал по параметрам x_2, x_3 . Однако, хотя это и разумно с точки зрения статистической теории распознавания образов, отнюдь не ясно способна ли нервная система реализовать сложные решающие функции. Можно предположить более простое решение вопроса, когда по каждому из акустических параметров принимаются первичные независимые решения, дальше сумма этих решений используется для выбора фонемы.

Были приведены эксперименты (4), в которых менялись два параметра, различающие *б* и *м*. Одним из них было изменение основной частоты при переходе от согласного к гласному, другим — спектральная характеристика согласного. Спектр согласного был упрощенным: он состоял из первой гармоники, равной 250 гц и четвертой гармоники, равной 1000 гц. Уровень интенсивности четвертой гармоники был переменным. В одной серии измерений экспериментатор устанавливал различные значения I_4 уровня четвертой гармоники в спектре согласного, испытуемый изменял значения основной частоты гласного, ища границу между *б* и *м*. В другой серии экспериментатор задавал значения основной частоты гласного, испытуемый менял уровень четвертой гармоники, ища ту же границу.

Полученные результаты приведены на рис. 6. По оси абсцисс отложен уровень интенсивности четвертой гармоники в спектре согласного, по оси ординат — величина изменения основной частоты при переходе от согласного к гласному. Ожидалось, что граница между фонемами будет функцией обоих параметров и будет представлена некоторой кривой, расположенной под углом к координатным осям. Как можно видеть, это предположение не оправдалось. Вместо ожидаемой одной границы, разделяющей плоскость

на две фонемные области, оказались две независимых границы, каждая по своему параметру; и пространство разделилось на четыре области.

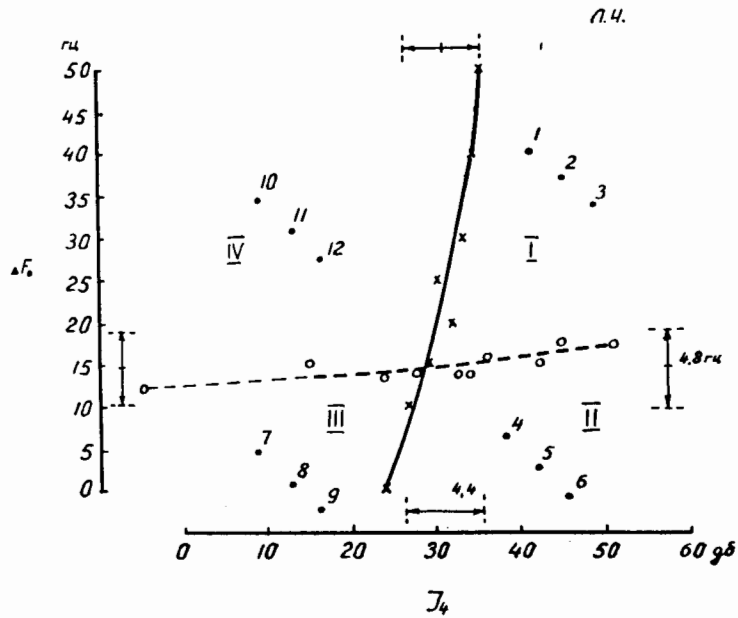


Рис. 6.

Картина оказалась аналогичной у всех трех испытуемых, только положение границ было несколько разным. На рис. 7 приведены данные другого испытуемого.

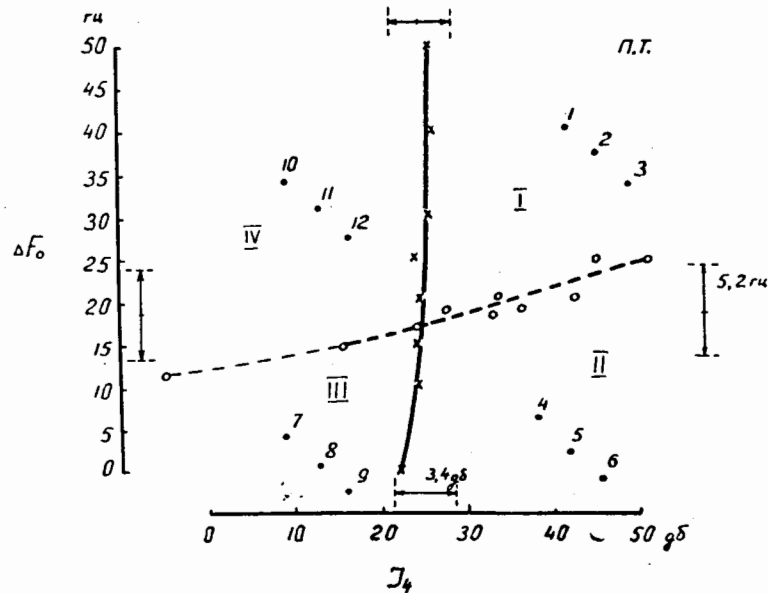


Рис. 7.

Полученные данные свидетельствуют, что по каждому параметру есть свой порог, и решение о том, по какую сторону от порога находится сигнал по одному параметру, принимается независимо от того, каковы значения сигнала по второму параметру. Иначе говоря, мы как будто бы имеем дело с системой, состоящей из параллельных каналов с фиксированными решающими правилами в каждом из каналов. Однако, оказывается также, что окончательная фонемная интерпретация стимулов не является постоянной. Стимулы из областей 1 и 3 в одной серии измерений воспринимались как б, а в другой — как м. Когда испытуемый менял уровень интенсивности четвертой гармоники и сравнивал между собой сигналы областей 4 и 1,

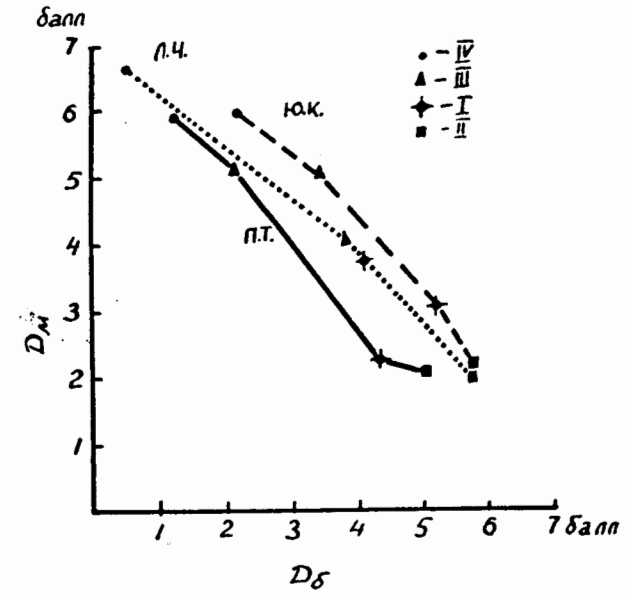


Рис. 8.

он трактовал стимулы области 1 как м. Когда он менял величину скачка основной частоты и сравнивал сигналы областей 2 и 1, он трактовал стимулы области 1 как б. Если мы имеем дело со схемой, принимающей параллельные независимые решения по отдельным акустическим параметрам, возможность различной интерпретации стимулов областей 1 и 3 представляется довольно логичной. Это действительно сигналы, которые по одному признаку являются фонемой б, а по другому — фонемой м.

Здесь однако возникает весьма интересный вопрос. Являются ли различные признаки равноценными, или один имеет больший вес, чем другие при окончательной фонемной интерпретации стимула? Ответить на этот вопрос методом исследования фонемных границ уже нельзя. Была сделана

попытка применить для этой цели метод шкалирования субъективных расстояний между стимулом и фонемой. Значения выбранных 12 стимулов показаны точками на рис. 7. Стимулы были записаны на пленку в случайной последовательности. Они предъявлялись тем же трем испытуемым. В одной серии прослушиваний испытуемые определяли для каждого стимула его расстояние в баллах от фонемы *б*, в другой серии — от фонемы *м*.

Полученные результаты приведены на рис. 8. Даны средние расстояния в баллах для стимулов каждой из четырех областей. По оси абсцисс отложено расстояние от фонемы *б*, по оси ординат — от *м*. Можно видеть, что стимулы области 4 оцениваются как близкие к *б* и далекие от *м*. Стимулы области 2 близки к *м* и далеки от *б*. Стимулы областей 1 и 3 занимают промежуточное положение, однако стимулы области 1 ближе к *м*, а стимулы области 3 ближе к *б*. Это уже говорит о том, что вес спектрального признака согласного больше веса признака скачка основной частоты.

Для количественного вычисления весов было сделано предположение, что субъективное расстояние между стимулом и фонемой пропорционально разности между некоторым постоянным числом, представляющим идеальную фонему, и суммой весов положительных признаков данной фонемы, присутствующих в стимуле. Оказалось, что относительные веса признаков довольно сильно различались у наших трех испытуемых. Так, у одного испытуемого вес признака скачка основной частоты составил примерно 0,80 от веса спектрального признака, тогда как у другого испытуемого он составил всего 0,16 от веса спектрального признака.

Настоящие данные слишком ограничены, чтобы на их основании можно было делать какие либо обобщения относительно истинной организации процесса восприятия. Однако, я все же рискну высказать некоторую гипотезу. Она состоит в том, что процесс принятия фонемных решений состоит из двух этапов. На первом этапе принимаются решения об отдельных полезных акустических признаках, устанавливается, в какой области значений по каждому из параметров находится сигнал. На этом этапе распознающая система действует как система с постоянными характеристиками: и способ измерения параметра и решающие правила фиксированы. На следующем этапе происходит суммирование первичных решений, помноженных на соответствующие веса, и сравнение их с некоторыми порогами. Логично предположить, что значения весов устанавливаются в процессе обучения, а значения порогов регулируются текущим образом в зависимости от контекста и других априорных сведений о том, какая фонема должна быть на данном месте. Такая гипотеза, напрашивающаяся на основании полученных данных, представляется достаточно физиологичной, т. к. соответствующая ей модель может быть довольно просто реализована с помощью нервной сети. Если эта модель соответствует истине, метод поиска границ годится для исследования первого этапа распознавания: вы-

явления и уточнения природы измеряемых слухом признаков сигнала и применяемых к ним решающих правил. Для исследования второго этапа необходимы другие методы.

ЛИТЕРАТУРА

1. Fant, G.: The nature of distinctive features. *STL-QPSR* 4/1966, 1—14.
2. Chistovich, L.—Fant, G.—de Serpa-Leitao, A.: Mimicking and perception of synthetic vowels. Part II. *STL-QPSR* 3/1966, 1—3.
3. Чистович Л. А.: Изменение основной частоты голоса как различительный признак согласных. *Акустический журн.* В печати.
4. Чистович Л. А.: О процедуре распознавания фонем человеком. *Вопросы психологии.* В печати.

DISCUSSION

Fischer-Jørgensen:

I find Dr. Chistovich's experiments very interesting. I only want to emphasize (and I do not think I really disagree with her on this point) that what is found by this method is not a *phoneme* boundary, but a boundary between two areas in the phonetic space which in this particular environment and in this particular situation are conceived as manifestations of two different phonemes. As the boundaries are very often different for different environments, we shall not expect to be able to find clear boundaries between areas covering all manifestations of given phonemes or distinctive features, but rather cases of complex overlappings. This was shown thirty years ago for German vowel quantity by E. Zwirner.

Lindner:

Die vorgetragenen Untersuchungen über die Bestimmung der Grenzen zwischen Klassen von Lauten decken sich im wesentlichen mit Untersuchungen, die in den letzten beiden Jahren an unserem Institut durchgeführt worden sind. Dabei zeigte es sich, daß die Grenzen zwischen Vokalklassen von der Herkunft des Perzipienten und seiner Zugehörigkeit zu einer Dialektgemeinschaft abhängig sind. Neben objektiv bestimmbareren Parametern des akustischen Signals müssen auch subjektiv bedingte Parameter berücksichtigt werden.

Tillmann:

I wonder whether decision theory by its purely statistical nature could help us much in finding a model of speech perception, even a neurological one. Decision theoretic work on pattern recognition has also shown empirically what is logically to be proved: that decision theory is blind with regard to certain kinds of structure.

Fujisaki:

I should like to express my respect for the extensive investigation which Dr. Chistovich has just reported, particularly because I myself have been conducting research on the perceptual boundaries between Japanese vowels. May I make two comments:

(I) I believe that the experiments on the boundary between (ba) and (ma) would have been more meaningful if the two parameters were varied not independently, but in accordance with relationships found in natural speech.

(II) I should like to ask if Dr. Chistovich would present a physiological model which would support the particular shape of the boundary obtained in the F_0 — ΔF_0 plane.

Загоруйко:

Я хотел бы выступить в защиту статистической теории распознавания образов. В терминах этой теории можно объяснить все те интересные экспериментальные данные, которые получены в работе Л. А. Чистович. В частности, можно легко показать, что механизм принятия решения на основании порогового измерения отдельных признаков проще (экономичнее), чем с помощью многомерной разделяющей гиперповерхности.

DES INDICES ACOUSTIQUES AUX TRAITS PERTINENTS

PIERRE DELATTRE*

INTRODUCTION

Les recherches sur la nature des traits pertinents par lesquels les phonèmes d'une langue donnée se distinguent les uns des autres doivent commencer au niveau acoustique (physique) plutôt qu'articulatoire, nous semble-t-il, parce que le fonctionnement des traits est un phénomène de perception, et que, du moins dans l'état présent des facilités de recherche, seuls les corrélatifs acoustiques peuvent être soumis au test perceptuel. En effet, grâce aux techniques expérimentales de synthèse établies par les Laboratoires Haskins, la manipulation de formes acoustiques appropriées, au moyen d'un synthétiseur de parole, permet d'isoler, de modifier, de combiner, chacun des éléments acoustiques (physiques) qui contribue à la perception du lieu ou du mode d'articulation d'un phonème donné, permet d'en faire varier une seule dimension à la fois et de tester à l'oreille les effets de ces variations; cette manipulation des formes synthétiques permet même d'établir quels sont les éléments acoustiques qui ne contribuent pas à la perception de tel ou tel trait pertinent ou qui n'y contribuent que dans une certaine mesure. Or il n'est pas possible d'isoler, de modifier, de disséquer, de combiner les éléments d'un trait *articulatoire* et de tester séparément à l'oreille l'effet des variations de chacun des éléments. Il est possible par contre de décrire avec un assez haut degré d'exactitude les mouvements articulatoires élémentaires qui correspondent à des indices acoustiques élémentaires une fois que ceux-ci ont été bien établis, limités et spécifiés par les tests perceptuels. Si donc on ne parvient pas à isoler et spécifier les corrélatifs articulatoires des traits pertinents des phonèmes directement, on peut du moins le faire indirectement par l'intermédiaire des indices acoustiques obtenus par les tests de perception.

Pour illustrer l'état présent des connaissances acoustiques sur la parole, et surtout les avancées récentes dans le domaine des modes d'articulation (les lieux d'articulation ont déjà fait l'objet de nombreuses études), nous nous proposons de présenter ici le tableau complet de la structure acoustique des consonnes intervocaliques d'une langue, en tant que système. Nous avons choisi le français plutôt que l'anglais (pourtant mieux étudié) parce que les problèmes d'acoustique qu'on y rencontre sont d'une nature plus universelle.

* Université de Californie (Santa Barbara).

La Figure 1 est un tableau structural des formes spectrographiques des consonnes françaises entre voyelles [ε]. L'arrangement des formes en sept colonnes et en huit rangées est entièrement basé, non sur les traits articulatoires mais sur les traits acoustiques et le genre de perception qu'ils caractérisent. Ainsi les consonnes d'une

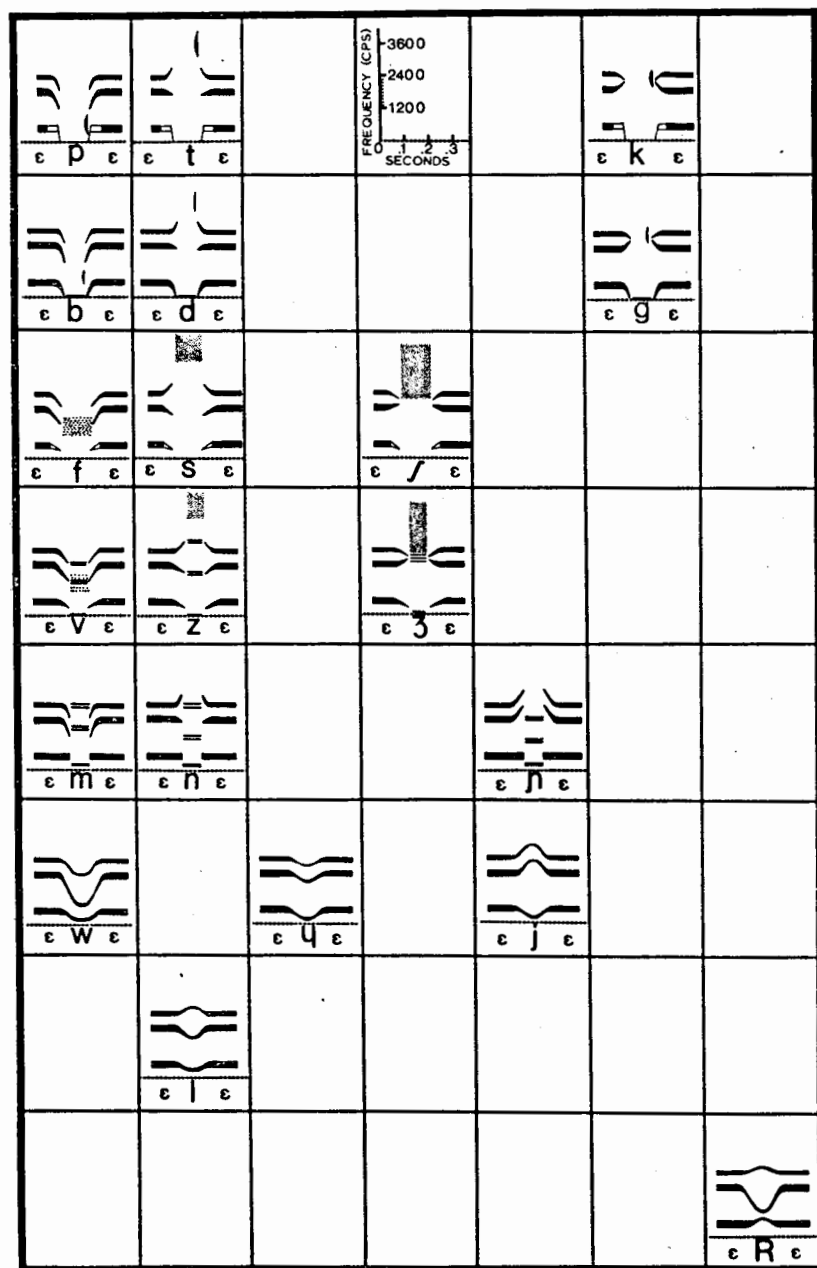


Fig. 1.

même colonne ont en commun des indices *acoustiques* qui ont quelque rapport avec la perception des lieux d'articulation, et celles d'une même rangée ont en commun des indices *acoustiques* qui ont quelque rapport avec la perception des modes d'articulation.

DÉMONSTRATION DE PAROLE ARTIFICIELLE

[Ici, les enregistrements au magnétophone des formes acoustiques de la Figure 1 ont été entendus sous trois aspects différents: a) transformées en son voisé par un synthétiseur du type „playback“ à 50 bandes de fréquence, sans variations de la fréquence fondamentale; b) transformées en son voisé par un synthétiseur du type „voback“ à 18 bandes de fréquence, avec variations de fréquence fondamentale; c) transformées en son chuchoté par le même voback n'utilisant que la source de bruit. Les voyelles chuchotées font parfaitement entendre la distinction dite „sourdes/sonores“, ce qui montre bien le poids des indices acoustiques autres que la fréquence fondamentale produite par la vibration des cordes vocales. (Ces indices seront démontrés plus loin). D'ailleurs, l'intelligibilité des syllabes chuchotées, dont les formants s'obtiennent par filtrage de bruit blanc, est tout aussi claire que celle des syllabes voisées. Cela démontre la prépondérance du „formant“ dans la perception: les sons vocaliques ne sont pas caractérisés par des harmoniques mais par des formants (bandes étroites de son filtrées par la résonance sélective du tractus vocalique), que la source au larynx soit composée de son périodique ou de son turbulent.]

LES INDICES DE LIEU D'ARTICULATION

Sur la Figure 1, on peut voir que toutes les formes spectrographiques d'une même colonne ont en commun deux indices acoustiques de transition (deuxième et troisième formants), et que de plus toutes les explosives et fricatives d'une même colonne ont en commun la fréquence du bruit, que ce soit un bruit d'explosion ou un bruit de friction.

Première colonne (en partant de la gauche)

Les transitions du deuxième formant visent toutes un locus de très basse fréquence (700 cps environ). Les transitions du troisième formant visent toutes un locus de 2000 cps environ. Les bruits, sous forme d'explosion pour (p) et (b) ou de friction pour (f) et (v), ont tous une fréquence très basse (1000 cps environ).

[On donne le nom de *locus* au point de convergence de toutes les transitions de formant qui sont reliées à la perception d'un même lieu d'articulation consonantique ou de tout autre trait pertinent, quelle que soit la voyelle qui précède ou qui suit. Il est pratique de considérer le locus comme un point virtuel placé à quelque 5 cs des têtes de transition dans le cas des explosives et des fricatives; et comme un point réel dans tous les autres cas. La théorie du locus décrit la direction que les transitions de formant doivent prendre pour contribuer à la perception d'un lieu d'articulation ou de tout autre trait pertinent consonantique dans une langue donnée. Etant fondée non

sur l'observation seule des spectrogrammes mais essentiellement sur des tests perceptuels de variables synthétiques, cette théorie permet de juger, par extension, les conditions dans lesquelles telle transition de formant ne peut pas contribuer à la perception de tel trait pertinent, donc de déterminer sur les spectrogrammes de parole naturelle quelle est la contribution perceptuelle de chaque transition de formant. Ainsi lorsque la coarticulation empêche une transition de se diriger vers son locus, comme c'est le cas pour [kɔ, ko, ku], les tests de perception indiquent clairement que cette transition ne contribue pas à la perception du trait pertinent voulu et qu'un autre indice s'est substitué à la transition inefficace. Les tests perceptuels de parole synthétique ont d'ailleurs été confirmés par des expériences de parole naturelle faites par découpage de bandes magnétiques.]

Deuxième colonne

Le locus des transitions du deuxième formant est à environ 1700 cps — compromis économique possible entre le locus de (t, d, n) à 1800 cps et celui de (s, z, l) à 1600 cps. Le locus des transitions du troisième formant est à environ 2700 cps. Les bruits sont relativement hauts: vers 4000 cps pour les explosions de (t, d), en dessus de 3500 cps pour les frictions de (s, z). Dans cette colonne, en dehors du (l), le locus du troisième formant a plus de pouvoir perceptuel que dans les autres colonnes. Le locus du troisième formant est légèrement moins élevé et moins efficace pour (l) que pour les autres consonnes de la même colonne, mais il ne semble pas nécessaire d'ouvrir une colonne particulière pour cette résonnante. Bien que latéral, un (l) français est assez antérieur pour s'identifier *acoustiquement* avec les consonnes dentales. Ce ne serait pas le cas d'un (l) anglais.

Troisième colonne

Pour la semi-voyelle (ɣ), il a fallu ouvrir une colonne spéciale. Bien qu'arrondie du point de vue labial, son locus de deuxième formant coïncide plutôt avec celui des dentales (1700 cps), lesquelles ont généralement les lèvres écartées, et bien que palatale du point de vue lingual, son locus de troisième formant (très efficace) coïncide plutôt avec celui des alvéolaires (2200 cps). Cela nous montre que les traits articulatoires n'agissent pas sur les traits acoustiques d'une manière indépendante mais d'une manière combinée. On y trouve aussi un rappel du fait que le seul terme „labial“ ne suffirait pas à définir le commun dénominateur articulatoire de la première colonne. La cinéradiographie montre en effet que les consonnes de la première colonne ont toutes, en plus de leur labialisation, une cavité post-labiale relativement grande, soit par recul, soit par abaissement de la langue.

Quatrième colonne

Le locus des transitions du deuxième formant est à environ 2000 cps, celui des transitions du troisième formant à environ 2200 cps, et le bruit de friction à une fréquence intermédiaire, au dessus de 2000 cps. Le fait que la différence entre les

traits acoustiques de la quatrième colonne et ceux de la deuxième sont si marqués montre combien le léger recul du point d'articulation, de pré-alvéolaire à post-alvéolaire, peut être critique.

Cinquième colonne

Le locus des transitions du deuxième formant est à environ 3000 cps, et celui des transitions du troisième formant à environ 3500 cps. Nous remarquons que la consonne dite „palatalisée“ (ɲ) partage les traits acoustiques de lieu d'articulation avec la consonne palatale (j), ce que font également les autres consonnes dorso-palatalisées (tʃ, dʒ, kʃ, gʃ). Le terme „palatalisé“ ne désigne donc pas un mode d'articulation mais un lieu, et est équivalent de dorso-palatal.

Sixième colonne

Le locus des transitions du deuxième formant est à environ 3000 cps et celui des transitions du troisième formant à environ 2200 cps. Le bruit d'explosion est à une fréquence intermédiaire qui suit la tête de transition du deuxième formant et varie donc considérablement selon la fréquence du deuxième formant de la voyelle qui suit.

Un problème assez grave se présente ici du fait que plus la voyelle qui suit (k, g) est arrondie et postérieure, plus il est difficile à la transition du deuxième formant de se diriger vers le très haut locus de 3000 cps et par conséquent de contribuer à la perception du lieu d'articulation palatovélaire. Nous avons là une excellente illustration des méfaits de la coarticulation: l'arrondissement des lèvres et le recul de la langue pour une voyelle arrondie postérieure, comme (o) agrandit et ferme la cavité antérieure ce qui donne une note de résonance très basse à la tête de transition. Par bonheur, plus la voyelle est arrondie et postérieure, plus l'explosion de (k, g) a de pouvoir perceptuel. C'est donc l'indice d'explosion qui se substitue à celui de transition.

Septième colonne

De même que pour (ɣ), il a fallu créer une colonne spéciale pour (R), la forme acoustique des deuxième et troisième formants présentant une combinaison unique — la transition du deuxième formant s'apparente à celle de (w) et celle du troisième à celle de (l).

On remarque d'emblée que les relations acousticoarticulatoires qui prévalaient dans le cas de (w) et (l) n'ont plus cours ici — la consonne (R) n'est ni labio-vélaire ni apico-dentale, elle est pharyngale et pendant son articulation les lèvres sont le plus souvent écartées. Cela nous montre une fois de plus que ce n'est pas la labialisation en soi qui fait descendre la transition du deuxième formant mais toute condition articulatoire passible de faire résonner la cavité antérieure à une note basse: ici la note de tête du deuxième formant est basse à cause de l'immensité de la cavité antérieure (elle va du haut-pharynx aux lèvres); cette note serait d'ailleurs plus

basse même pour (*R*) que pour (*w*) si la grandeur de la cavité antérieure n'était compensée par l'écartement des lèvres.

Il nous faut dire un mot du premier formant de (*R*). On voit sur la Figure 1 que son locus est très haut (environ 700 cps), ce qui indique, d'après la théorie acoustique, que la constriction de cette consonne est dans la moitié postérieure du tractus vocalique, c'est-à-dire dans le pharynx. On pourrait donc dire que le haut locus du premier formant est un indice de lieu d'articulation. Il nous paraît cependant préférable de le ranger parmi les indices de *mode* d'articulation. Deux raisons suffiront à notre argument. a) Le haut locus du premier formant se trouve dans toute la catégorie des consonnes pharyngales, que leur constriction se trouve dans le haut pharynx, ou le bas pharynx, comme c'est le cas dans les langues arabes, ou qu'elle soit intermédiaire comme en français ou en allemand. b) Partout ailleurs que pour les (*R*), la transition du premier formant est exclusivement un indice de mode — c'est en partie la hauteur du locus de premier formant qui distingue, par exemple, les semi-voyelles des latérales, ou les occlusives des fricatives. Structuralement, donc, nous préférons considérer la hauteur du locus de premier formant des (*R*) français comme un indice de mode d'articulation.

LES INDICES DE MODE D'ARTICULATION

Examinons maintenant les rangées de la Figure 1. Ici il sera question des transitions de *premier* formant plutôt que de deuxième ou de troisième; de la *durée* des bruits plutôt que de leur fréquence; du *tempo* des transitions supérieures plutôt que de leur locus; du *caractère* des joints périodiques qui résonnent pendant la tenue d'une constriction plutôt que de leur fréquence.

Première et deuxième rangées

Voyons d'abord ce que ces deux rangées ont en commun.

1. La forme et le tempo de leurs transitions peuvent s'appeler „directe-rapide“. Les transitions des explosives, comme celles des fricatives et des nasales, sont *directes* par opposition aux transitions des résonnantes (nasales, semi-voyelles et liquides) qui sont *inverses* (leur courbe change de sens, comme celle d'un *S*); elles sont rapides, comme les transitions des nasales, par opposition à celles des fricatives et des glissantes (semi-voyelles et liquides). Ainsi les caractéristiques de forme et de tempo rangent les transitions en quatre types: directe-rapide pour les explosives, directe-lent pour les fricatives, inverse-rapide pour les nasales, et inverse-lent pour les glissantes.

2. Les transitions du premier formant visent la fréquence la plus basse possible, ce qui fait dire, sans doute arbitrairement, que leur locus est à zéro. Le locus de premier formant des fricatives, des nasales et des semi-voyelles est plus haut (250 cps), celui des latérales encore plus haut (400 cps) et celui des pharyngales le plus haut de tous (700 cps).

3. Les consonnes des deux premières rangées ont du bruit dans la tenue. Elles partagent cet indice avec les fricatives, mais pas avec les résonnantes, qui, elles,

sont privées de toute turbulence. D'autre part, les explosives n'ont pas de joint bas. Elles ne partagent cette absence de joint qu'avec les fricatives, la présence du joint bas étant l'indice le plus caractéristique des résonnantes (nasales, semi-voyelles et liquides). On peut définir le joint bas comme un formant de tenue, donc de faible intensité, qui relie les transitions implosives et explosives de premier formant. L'opposition (mutuellement exclusive) entre les bruits et les joints bas semble capitale parmi les indices consonantiques. Elle explique pour la première fois la différence essentielle entre les fricatives sonores et les semi-voyelles et liquides. (Étant corrélative de l'opposition transition directe/ transition inverse, elle peut être considérée comme redondante si l'on sépare, en ce qui concerne les transitions, la notion de forme de celle de tempo.)

4. Le bruit des explosives est court, par opposition à celui des fricatives qui occupe toute la tenue. C'est là évidemment l'indice principal pour distinguer les fricatives des explosives. Mais ce n'est pas le seul — le tempo de toutes les transitions et le locus des transitions de premier formant contribuent aussi à les distinguer.

5. Les quatre catégories d'indices que nous venons de décrire s'appliquent aussi bien à la seconde qu'à la première rangée de la Figure 1. Examinons maintenant les indices qui distinguent la première rangée de la seconde. Pour les explosives, les expériences de synthèse ont permis d'isoler six différents indices de cette opposition: toutes les formes spectrographiques de la première rangée ont une plus longue tenue que celles de la deuxième rangée; elles ont aussi une voyelle plus courte, une réduction (cutback) de la transition du premier formant, un bruit plus intense, une absence de fondamental, et enfin un peu d'aspiration si le sujet parlant est un aristocrate du Faubourg Saint-Germain.

Troisième et quatrième rangées

Examinons d'abord les indices acoustiques qui sont communs à ces deux rangées.

1. Leurs transitions sont directes-lentes — directes, comme celles des explosives, mais lentes, contrairement à celles des explosives.

2. Le locus du premier formant des fricatives est plus élevé que celui des explosives. On peut soit le considérer comme virtuel à environ 250 cps, soit le considérer comme réel à environ 350 cps.

3. Les fricatives ont un bruit de tenue et n'ont pas de joint bas. Nous l'avons déjà dit, c'est sur cette opposition d'indices acoustiques que repose principalement la distinction entre fricatives sonores et résonnantes.

4. Les indices acoustiques de bruit sont dits longs pour indiquer qu'ils occupent toute la durée de la tenue.

5. Ainsi les troisième et quatrième rangées ont quatre types d'indices acoustiques en commun. Voyons maintenant ce qui les oppose. Pour les fricatives, les expériences de synthèse ont permis d'isoler six indices acoustiques, dont cinq ont déjà été mentionnés pour les explosives. Les formes de la quatrième rangée ont des joints périodiques qui relient les transitions implosives aux transitions explosives de deuxième et

troisième formants. Ces joints ou murmures oraux de faible intensité n'apparaissent pas dans la troisième rangée. Cette rangée est de plus caractérisée par un bruit plus long, une voyelle précédente plus courte, une réduction de la transition du premier formant, un bruit de friction plus intense, et l'absence de fondamental.

Cinquième rangée

1. Les transitions de formant de la cinquième rangée sont dites rapides-inverses. Elles sont aussi rapides que celles des explosives dans les deuxième et troisième transitions et plus rapides encore dans la première transition. Quant à la forme, leur inversion est plus angulaire que celle des glissantes mais c'est tout de même une inversion.

2. En ce qui concerne le locus de premier formant des consonnes qui ont un joint bas, on conviendra de le spécifier par la fréquence du joint bas en disant que c'est un locus réel (non virtuel). D'après cela, le locus de premier formant des nasales est à 250 cps.

3. Toutes les consonnes nasales ont un joint bas à environ 250 cps et sont par conséquent privées de bruit (le joint bas et le bruit sont mutuellement exclusifs).

4. Enfin les consonnes nasales ont un indice acoustique qu'elles ne partagent avec aucune autre classe de consonnes: la discontinuité. Quand les autres consonnes ont un joint (haut ou bas), ce joint est en continuité avec les transitions implosives et explosives qu'il relie. Dans le cas des nasales, tous les joints (hauts et bas) montrent un certain degré de discontinuité.

Sixième, septième, et huitième rangées

Les semi-voyelles et les liquides de ces trois rangées partagent tous les indices acoustiques de mode d'articulation sauf la fréquence du locus de premier formant, qui, par convention, fournit automatiquement la fréquence du joint bas. Ces trois rangées de consonnes ayant beaucoup en commun du point de vue acoustique, nous avons cru utile de leur donner le nom commun de „glissantes“ (glides).

1. Toutes les glissantes ont des transitions de formant inverses-lentes — caractéristique qui leur est unique.

2. Elles ont toutes un joint bas (beaucoup plus indispensable que les joints hauts, dans la synthèse) et par conséquent sont privées de tout bruit.

3. Leurs joints sont tous en continuité avec les transitions implosives et explosives qui précèdent et suivent.

4. Elles se répartissent en trois groupes selon la fréquence du locus du premier formant: les semi-voyelles de la première rangée ont un locus de premier formant à environ 250 cps, comme les nasales et les fricatives; la latérale de la septième rangée a son locus de premier formant à environ 400 cps; et la pharyngale de la huitième rangée a son locus de premier formant à environ 700 cps. (Nous avons donné plus haut nos raisons pour traiter le haut locus de la consonne (*R*) comme un indice de mode d'articulation plutôt que de lieu.)

LES CORRÉLATIFS ARTICULATOIRES DE MODE D'ARTICULATION

Voici les traits articulatoires qui correspondent aux traits acoustiques qui ont pu être isolés. On suivra l'ordre de la Figure 2.

PLACE		MURMURE		CUES		A		B		C				
						Trans. 2		Trans. 3		Turb.				
						Locus at 700	Locus at 1700	Locus at 2000	Locus at 3000	Locus at 2700	Locus at 3500	p-f Low	k-s Mid	t-s High
		w	m	v	f	b	p	x	-	-	-	x	-	-
	l	n	z	s	d	t		-	x	-	-	-	x	-
	ç							-	x	-	-	-	-	-
			3	ç				-	-	x	-	-	-	x
	j	p						-	-	-	x	-	-	-
					g _a	k _a		-	-	x	-	x	-	-
					g _o	k _o		-	-	-	-	x	-	-
	R							x	-	-	-	-	x	-
-	-	-	-	x	x	x	x	Direct Transitions		(Reverse Transitions)				
-	-	-	x	-	-	x	x	Fast Transitions		(Slow Transitions)				
-	-	-	-	-	-	x	x	F ₁ Locus at 0		(F ₁ Locus not at 0)				
-	-	x	x	x	x	-	-	F ₁ Locus at 250		(F ₁ Locus not at 250)				
-	x	-	-	-	-	-	-	F ₁ Locus at 400		(F ₁ Locus not at 400)				
x	-	-	-	-	-	-	-	F ₁ Locus at 700		(F ₁ Locus not at 700)				
-	-	-	-	x	x	x	x	Turbulence		(No Turbulence)				
x	x	x	x	-	-	-	-	Low Periodic Link		(No Low Periodic Link)				
				-	-	x	x	Short Turbulence		(Long Turbulence)				
-	-	-	x	-				Discontinuous Links		(Continuous Links)				
x	x	x	x	x	-	-	-	High Periodic Links		(No High Per. Links)				
				x	-	x	-	Short Hold		(Long Hold)				
				x	-	x	-	Long Preceding Vowel		(Short Prec. Vowel)				
				x	-	x	-	No Cutback		(Cutback)				
				x	-	x	-	Weak Turbulence		(Strong Turbulence)				
				x	-	x	-	Voice Bar		(No Voice Bar)				
							x	No Aspiration		(Aspiration)				

Fig. 2.

Une transition directe correspond articulatoirement à une production de bruit à l'initiation du mouvement d'ouverture (explosives) ou juste avant (fricatives); une

transition inverse correspond à une absence de bruit à l'initiation du mouvement d'ouverture (nasales) ou avant (glissantes).

Une transition rapide correspond à une occlusion orale complète pendant la tenue (explosives et nasales); une transition lente correspond à une occlusion orale incomplète pendant la tenue (fricatives et glissantes).

Un locus de premier formant à zéro correspond à une interruption complète de la colonne d'air par une constriction dans la partie antérieure du tractus vocalique.

Un locus de premier formant à 250 cps correspond à une interruption centrale (non latérale) incomplète dans la moitié antérieure du tractus.

Un locus de premier formant à 400 cps correspond à une interruption latérale incomplète dans la moitié antérieure du tractus.

Un locus de premier formant à 700 cps correspond à une constriction centrale dans la moitié postérieure du tractus vocalique, c'est-à-dire dans le pharynx.

Un bruit sans joint bas pendant la tenue ou à son terme correspond à un souffle dont la pression s'exerce derrière une constriction étroite ou une occlusion complète et qui passe, soit pendant la tenue (fricatives), soit au terme de la tenue (explosives).

Un joint bas sans bruit pendant la tenue correspond à un souffle qui traverse une constriction orale ou vélique (voile du palais abaissé), laquelle n'est pas assez étroite pour ralentir l'écoulement du souffle et permettre à la pression d'air d'augmenter derrière la constriction. La constriction ne sert en somme qu'à séparer les cavités de résonance.

Un bruit court correspond à l'ouverture brusque d'une fermeture complète derrière laquelle la pression du souffle a été en augmentant.

Un bruit long (occupant toute la tenue) correspond à un souffle qui traverse de façon ininterrompue une constriction étroite.

Des joints de tenue discontinus (consonnes nasales) correspondent à la réunion et à la séparation des organes de constriction pendant que le voile du palais est abaissé. (Les joints sont discontinus même si la consonne nasale est précédée et/ou suivie d'une voyelle *nasale*.) Les deux instants de discontinuité (aux jonctions des joints avec le terme des transitions implosives et le départ des transitions explosives) correspondent aux instants où, a) l'occlusion buccale se fermant, le souffle est brusquement dérivé vers le nez, et b) l'occlusion buccale se rouvrant, le souffle est brusquement dérivé vers la bouche. A ces deux instants le système de résonance du tractus change radicalement.

Des joints hauts sans joint bas (comme pour les fricatives sonores) correspondent à un état de résonance du tractus qui permet à la seule cavité antérieure (entre la constriction et les lèvres) de résonner effectivement. L'inefficacité de la cavité postérieure vient de ce que la glotte fait double emploi: elle s'ouvre à l'arrière pour laisser passer le souffle librement et se ferme à l'avant pour vibrer périodiquement. L'ouverture à l'arrière fait communiquer le pharynx avec la trachée, ce qui abaisse et réduit sa fréquence de résonance. Cette inefficacité doit aussi s'attribuer au fait

que la constriction des fricatives — même des fricatives sonores — est trop étroite pour laisser passer les vibrations de la cavité postérieure.

Une longue tenue acoustique correspond à une longue tenue articuloire des organes au degré d'aperture de cette consonne.

Une longue voyelle précédente correspond à une longue tenue des organes au degré d'aperture de cette voyelle.

La réduction des transitions de premier formant (cutback) correspond au degré exceptionnel de pression musculaire qui s'exerce lorsque les organes se joignent ou se rapprochent étroitement pour une consonne dite sourde.

Un bruit fort correspond à un haut degré de pression derrière une constriction tendue.

L'aspiration correspond à un délai dans la mise en vibration des cordes vocales après l'ouverture brusque de l'occlusion, délai pendant lequel la glotte est ouverte et le débit de souffle est violent (sept fois plus grand que pour une voyelle chuchotée) ce qui ne permet qu'à la cavité antérieure de résonner efficacement (pour l'aspiration, comme pour le *(h)*, le premier formant vocalique est fort atténué, tandis que pour la voyelle chuchotée, il est renforcé).

SOMMAIRE DES CORRÉLATIFS ACOUSTIQUES

La Figure 2 résume les résultats des pages précédentes.

Le carré central présente le même arrangement de consonnes que le tableau structural des formes spectrographiques de la Figure 1, sauf qu'il y est réservé deux rangées aux occlusives vélares (*(k)* et *(g)*) parce que leurs spécifications acoustiques sont différentes selon qu'elles sont suivies de voyelles écartées (symbolisées par *k_eg_e*) ou de voyelles arrondies (symbolisées par *k_og_o*). Nous avons en effet vu plus haut que devant les voyelles arrondies-postérieures, les transitions du deuxième formant ne peuvent pas se diriger vers le locus de 3000 cps et que c'est l'explosion, au lieu de la transition, qui fait percevoir le lieu d'articulation vélaire. En somme, acoustiquement parlant, on identifie les occlusives vélares par deux types d'allophones extrêmement différents l'un de l'autre.

L'arrangement du carré de consonnes montre que l'analyse acoustique répartit les consonnes françaises en sept catégories de lieu d'articulation (rangées sur la Figure 2) et huit catégories de mode d'articulation (colonnes sur la Figure 2). Mais il en est tout autrement des indices acoustiques mêmes qui permettent l'identification linguistique de ces consonnes.

D'après la Figure 2, l'identification linguistique des consonnes utilise au moins 11 indices acoustiques de lieu (extensions du carré vers le haut) et 17 indices acoustiques de mode (extension du carré vers le bas). Les matrices de signes „plus“ et „moins“ indiquent pour chaque consonne le choix qui s'opère parmi les indices de lieu et de mode.

Les 11 indices de lieu se rangent en trois catégories: A. La direction de la transition du deuxième formant, spécifiée par un locus. B. La direction de la transition du

troisième formant, également spécifiée par un locus. C. La fréquence du bruit (seulement les explosives et les fricatives), lequel peut se réaliser par une explosion ou une friction. En français, la largeur de bande et l'intensité ne sont pas pertinentes.

Les 17 indices de mode se rangent en six catégories:

1. La forme et le tempo des transitions.
2. Le locus de la transition de premier formant.
3. La présence de bruit, contre la présence de joint bas.
4. La longueur du bruit (s'il y en a). Cela remplace l'opposition conventionnelle: interrompu/non-interrompu.
5. La continuité contre la discontinuité des joints (s'il y en a).
6. Les facteurs de durée et d'intensité, communément appelés „sourd“ contre „sonore.“

Les cases remplies des matrices révèlent les faits suivants:

a) Deux consonnes ayant le même mode d'articulation ne se distinguent jamais l'une de l'autre par moins de deux indices de lieu ou par plus de trois.

b) Deux consonnes ayant le même lieu d'articulation ne se distinguent jamais l'une de l'autre par moins de deux indices de mode ou par plus de 11 (par plus de 5 si les indices de voisement sont comptés séparément).

Les conditions (a) et (b) montrent bien la complexité des corrélatifs objectifs des traits pertinents: Il n'y a pas d'indice acoustique qui soit distinctif par lui-même, c'est-à-dire qui puisse convertir une consonne en une autre par substitution; il faut au moins deux indices et jusqu'à onze pour produire une distinction de sens. Le corrélatif acoustique d'un trait pertinent (appelons-le „indice distinctif“) se présente donc sous la forme d'un faisceau d'indices contributifs, en français, et il y a de bonnes raisons de croire que le corrélatif articulatoire d'un trait pertinent (appelons-le „articulation distinctive“) n'est pas moins complexe.

[Il a été donné en conclusion une démonstration sonore de la contribution perceptuelle à l'identification des consonnes françaises, de chacun des 11 indices de lieu d'articulation et des 17 indices de mode, au moyen de diapositives des formes spectrographiques et d'enregistrements magnétiques de parole artificielle.]

Disons pour terminer que si les traits pertinents sont des signaux perceptuels qu'on ne peut pressentir qu'indirectement à travers leurs corrélatifs acoustiques et articulatoires, et que les corrélatifs articulatoires ne peuvent être spécifiés qu'une fois accompli l'isolement des corrélatifs acoustiques, il n'est peut-être pas possible de toucher les traits pertinents de plus près qu'en arrivant à une connaissance suffisante de ce qui est distinctif dans les signaux acoustiques. C'est à quoi nous avons visé ici.¹

¹ The research reported herein was performed pursuant to a contract with the United States Department of Health, Education and Welfare, Office of Education.

DISCUSSION

Achmanova:

The "vision", the functional, the *humanistic* approach Prof. Delattre has been speaking about *will* become useful and promising only when the *instrumental* studies have been, if not completed, at any rate carried far enough. I could not follow him in thinking that the minute instrumental "preliminaries" can be regarded as completed.

Tillmann:

Thank you for the many interesting details you presented. What I want to suggest is that you should restrict your results to artificial speech. Your cues and features are, perhaps, sufficient but not necessary properties of acoustic signals, perceived as speech. Nevertheless I am sure that the findings through analysis by synthesis are a good base to proceed from to the more sophisticated study of natural speech signals.

Wode:

Prof. Delattre has made it clear that he accepts a one-to-many relationship between acoustic cue and articulatory correlates. For establishing his acoustic cues he assumes a direct one-one correspondence between the auditory impression (satisfactory synthesis of the French consonants in terms of listener tests) and acoustic stimulus. This is in contradiction to much evidence that points to the fact that there is at least a many-to-one (if not a many-to-many) correspondence between acoustic cue and auditory impression. On the basis of this one would have to doubt the relevance of Prof. Delattre's conclusions concerning natural speech.

SOUND, FEATURES, AND PERCEPTION*

GUNNAR FANT**

THE SPEECH COMMUNICATION CHAIN

Speech communication may be considered as the transmission of information through a succession of stages within a speaker, a connecting medium, and a listener. Flow diagrams of this process can be elaborated in various forms depending on the detail of the analysis attempted and the aspects of the communication process on which the investigator focuses his descriptive efforts. The following tabulation of stages will be considered here.

- A. Production
 - (1) Intended meaning of message
 - (2) Message sentence form
 - (3) Neural production program
 - (4) Myodynamic activity
 - (5) Aerodynamic and acoustic processes
- B. Technical medium
 - (1) The acoustic speech wave emitted by the speaker
 - (2) Speech signal representation in various parts of a technical communication system
 - (3) The acoustic speech wave affecting the listener
- C. Perception
 - (1) Cochlear response
 - (2) Primary neural analysis
 - (3) Identification of phonetic elements
 - (4) Identification of sentence structure
 - (5) The message received

The terminal stages remain rather hypothetical in view of our limited insight in the organization of brain functions. Therefore, the formulation of the stages *A*(1), *A*(2) and *C*(4), *C*(5) above reflects our general concepts of successive levels of language structure rather than established neurological functions.

* Condensed version of the oral presentation at the 6th Int. Congr. of Phonetic Sciences. The detailed material including illustrations is to be found in the author's contribution to the forthcoming *Manual of Phonetics*, edited by B. Malmberg, North-Holland Publishing Co. [see Fant (1968)].

** Dept. of Speech Communication, Speech Transmission Laboratory, Royal Institute of Technology (KTH), 10044 Stockholm 70, Sweden.

Each stage is to be characterized by an inventory of specific signals specified by parameters which possess certain time and space characteristics that combine into patterns according to general rules and constraints. A major ambition is to derive rules for translating a representation on one stage to a corresponding representation on any other stage of the complete system. Stage *A*(4) which comprises the dynamics of the speech organs may accordingly be described by a set of time varying articulatory parameters. One of the primary aims of general phonetics and speech research is to derive the rules for translating from this articulatory stage to that of the speech wave *B*(1).

At the stage *A*(5) comprising the acoustic production processes the signal structure can be divided into source and filter categories and each of these may be considered at two substages. Thus the filter-function is initially represented by the vocal tract "area-function", i.e. its resonator dimensions, from which their sound shaping properties may be derived by acoustic theory. Similarly, the source has a primary aspect of mean pressures and flows characterizing the aerodynamics of the exhaled air whilst the superimposed periodic or random disturbances constitute the raw material of voiced and unvoiced sounds.

This model of a successivity of encoding stages that the speech message has to pass from the transmitter to the receiver through the entire speech communication chain cannot be quantitatively studied with the same rigor as for instance a telegraph communication system. The main purpose of the model is to serve as a frame for formulating research objectives and discussing descriptive theory whilst the application of a quantitative signal and information analysis generally is beyond our capacity.

One sometimes encounters statements proposing that the information rate is very low at higher brain centers and increases towards the periphery with a maximum at the speech wave. This reasoning suffers from a confusion of the message and signal aspects of the communication. Ideally, the message is the same at all stages and the rate of information flow thus the same everywhere. It is more valid to speak of an increasing redundancy in the sense that the signal structure gets more complex and utilizes a larger number of parallel pathways whilst the information remains the same. Even this statement is rather loose in view of our limited insight in the neurological levels.

At present it is not possible to accomplish anything like a complete description of signal structure at any stage with the exception of the acoustic speech wave where all details of the waveform may be sampled and studied. However, even if we cared to carry out a maximally detailed sampling it would not be worth the labor. Also, there exists an infinite variety of transformations for expressing one and the same fact by different parameters, i.e. by different descriptive systems. Thus, in spite of apparent visual differences a narrow-band spectrum contains essentially the same information as a broad-band spectrum.¹

¹ The signal data contained in a spectrogram are mathematically equivalent to that of an

We have to accept the limitation of any quantification being approximate only but we require that it shall preserve a maximum of message information with as simple a signal description as possible. The extent to which such "minimum redundancy" or "compact" descriptive systems can be worked out is first of all a matter of how well the investigator is acquainted with the stage and its constraints and how complex abstractions he is capable of introducing.

THE NATURE OF DISTINCTIVE FEATURES

Complete formant specifications of a piece of speech is of practical use for synthetic reproduction only and is too detailed for comparative phonetic studies. What we need is a phonetically oriented data sampling system that allows us to sample the speech wave less densely than at intervals of the inverse of the bandwidth. The segmentation theory outlined by Fant and Lindblom (1961) and Fant (1962A and B) is a starting point for developing such a system. Segmental boundaries are mainly derived from changes in the "manner of production" whilst the "place of production" determines a more continuously varying element of segment patterns, in the first place the *F*-pattern (F_1, F_2, F_3, F_4) reflecting the continuous movements of the speech articulators.

This system operates with a terminology of speech production categories that is in part identical to that of the distinctive feature system of Jakobson, Fant and Halle (1952). The main difference is that the distinctive feature system serves a phonemic minimal redundancy classification purpose whereas the segment classification of Fant and Lindblom accounts for any production category irrespective of its communicative significance and is thus more phonetically detailed.

It should be appreciated that distinctive features in the sense utilized by Jakobson, Fant and Halle (1952) primarily constitute a system for subdividing phonemes and other components of the message ensemble. A distinctive feature has certain correlates on each stage of the speech communication chain and these correlates are described in terms of various parameters and cues, e.g. formant locations. A distinctive feature is thus a unit of the message ensemble rather than a property of the signal ensemble. The term "distinction" or "minimal category" would have been more appropriate and might have led to less confusion concerning their nature and use.

oscillogram providing the phase information is retained in the spectral representation. Relative phases within the spectrum would mathematically account for one half of the information concerning the signal structure but they are of rather minor communicative importance. Spectrograms are not designed to preserve phase information which in effect reduces the "redundancy" of spectral specifications by a factor of two compared to oscillographic specifications.

A formant representation is more condensed than a harmonic representation since a small number of formants can have the same descriptive power as a large number of harmonics. This economy is generally gained at some reduction of the accuracy in signal analysis. However, the harmonic representation is more detailed only when the voice fundamental frequency is low. The information gained in a low F_0 harmonic spectrum concerns irregularities of the voice source rather than the more important properties of the vocal tract transfer function.

The distinctive features are not intended as absolute descriptors of spectrographic qualities. The production or speech wave correlate of any feature will differ somewhat with the particular context of simultaneous and subsequent features. The invariance is generally relative rather than absolute. For instance, an invariable cue of compactness is the higher F_1 of the compact phoneme compared to the non-compact phoneme in the same context irrespective of which minimal pairs are inspected.

The relation between phonemes or features on the message level to speech segments and parameters on the signal level is generally complex. One segment may contain information about several successive phonemes and a single phoneme is generally related to several successive segments of the speech signal. As a rule the number of segments determined according to the principle of Fant and Lindblom comes out to be larger than the number of phonemes in the utterance. However, this is not always the case since in less careful articulation one or several phonemes of the intended message turn out to be produced in an extremely reduced fashion or omitted altogether without affecting the intelligibility. In practice we do not measure the duration of phonemes in the speech spectrogram but we measure the duration of sound segments and other characteristics of the speech signal.

A feature classification system can thus retain more or less redundancy and it can be more or less representative of actual encoding dimensions of the speech signal. The system of Jakobson, Fant, and Halle is too condensed for practical purposes such as comparative phonetic studies and development of automatic speech recognition schemes. The strength and novelty of the system is that it attempts to break the barrier between phonology and phonetics, linking the theory of message signs with the theory of their physical realization through the concept of the speech signal as a multi-dimensional event.

However, the specific choice of units still remains a disputable compromise between the two aspects. The extreme minimum redundancy objectives inherent in phonemic analysis have been the guiding principle for the selection of features. Accordingly, these constitute a very condensed and handy set for transcription of speech messages. Most of the features represent conventional phonetic categories which undoubtedly have a physiological and psychological significance. In a more phonetically oriented solution, on the other hand, one should increase the number of features so as to avoid or at least reduce the number of features operating in both vowels and consonants. In search for independent units on the signal level as opposed to a linguistic message level one might have to include major allophones of a language. The underlying principle would be to search for an inventory of speech production categories at our disposal for programming the phonatory and articulatory events. EMG, cineradiography, and direct recordings of the dynamical patterning of speech articulation will be helpful tools for such studies.

The search for generative rules of speech production may be exemplified by the studies of some of my colleagues, Lindblom (1963), Öhman (1966, 1967), and Öhman and Lindqvist (1965), who have tackled the problems of formulating rules for predict-

ing vowel reduction, coarticulation, and intonation contours. Given a phonemic or allophonic unit of the assumed production inventory the corresponding speech wave realization may be thought of as the output of "black box" labelled production mechanism the input of which is the selected unit plus a set of other discrete units representing the immediate context of other simultaneous, preceding and following units, prosody included. By a consistent analysis in terms of such models it should be possible to reach a more profound insight in the actual inventory of independent signal categories.

The model of Öhman operates with separate sets of control signals for vowels and consonants and this principle is also followed by Borovičková and Maláč (1966). The frequent use of one and the same feature in vowels as well as in consonants of the Jakobson feature system cannot be supposed to reflect an actual sameness of neurological encoding. Thus it would not be hypothesized that one and the same neural motor command labelled compactness is triggered off in the production of a consonant (k) and a vowel (a).

As a consequence of the high degree of economy aimed at in the Jakobson system and the unavoidable pay off for this economy in terms of a reduced phonetic similarity of a feature in widely different contexts it is not advisable to scale the phonetic distance between two speech sounds in terms of the number of distinctive features by which the corresponding phonemes differ. An extreme example that I have elaborated on earlier, Fant (1966A), is that the last two phonemes of the word "wing" the (i) and the (ng) do not have any distinctive features in common as pointed out by Jakobson whereas the temporal contrast between the sound segments related to [i] and [$ŋ$] is minimal only. The place of articulation being the same and the consonant anticipated already by the nasalization of the [i] the transition from [i] to [$ŋ$] merely involves a closing gesture of the tongue towards the palate.

One weakness of the phonological feature system leading to this paradox is that the palatal articulation goes with compactness in the consonantal system and with noncompact acute sounds in the vowel system. However, from an abstract acoustic feature point of view the (uia) interrelation show some similarities with the (ptk) relations. The relation [p/t] is a good parallel to [u/i] acoustically and the analogous role of (k) and (a) can also be supported in spite of the articulatory sameness of (i) and (k). From a perception point of view this similarity is superficial. In my view vowels and consonants are perceived through separate "feature channels", if any.

SPEECH PERCEPTION

From the accumulated experience on speech perception and especially experiments with speech-like synthetic stimuli it is apparent that speech is perceived categorically, Liberman et al (1967). We respond phonemically and tend to identify phonemes and allophones in the first place even when we are asked to discriminate small variations in quality, Liberman et al (1957). According to Liberman et al (1963) this effect is pronounced with consonants, whereas vowels are not perceived catego-

rically. Stevens (1966) reports on categorical effects in vowel perception providing the vowel is embedded in a syllabic frame. This effect is interpreted by Stevens as an instance of a principle that all factors that contribute to make the stimulus or the general conditions of the experiment representative of actual speech condition the listener to perform in a "speech mode" characterized by his making message identifications rather than quality gradations, Stevens (1966). This effect is a result of the listener's language experience rather than a unique property of the acoustic signal, Liberman et al (1967), Stevens and House (1966).

The significance of the concept of distinctive features is quite apparent from perception experiments. However, some investigators have interpreted the term distinctive feature at its face value only and accordingly identified it with the concept of a single important parameter or a cue. This has caused some confusions and distrust in the principle of distinctive features. As already stressed a feature is a recurrent phonemic distinction within a language and a major purpose of perception research is to evaluate the physical parameters and cues which signal the distinctions and phonemes of a language. The term cue is the same as an important physical parameter but can also be more complex in the sense that certain parameters combine to a characteristic pattern.

The Haskins Laboratories' systematic studies of the perception of simple stylized formant patterns have contributed greatly to our knowledge of the perceptual significance of formant data. However, the potential risk when working with simplified synthetic stimuli is that they may become insufficient carriers of phonemic cues and that the conclusions drawn from such experiments will be valid for the particular synthesizer only and not for human speech. This was the cause of the somewhat pessimistic conclusions Liberman et al (1957) made concerning the ambiguity of acoustic data as opposed to articulatory data in a study of F_2 -locus as a cue for identifying (*d*) and (*g*). It can be shown that the syllables (*da*) and (*ga*) have approximately the same F_2 -transition but this ambiguity is resolved by combining F_3 , F_2 and the release burst into a single cue, Fant, Lindblom, and de Serpa-Leitão (1966), Fant (1968).

There remains much to be studied concerning the speech wave characteristics of phonemes and distinctions. A practical strategy is to start out with a detailed list of observable spectrographic pattern cues. In order to make a specification of contextual variants feasible it is advisable to present the data on each phoneme or feature in a few reference contents only and add contextual rules derived from studies of coarticulation, reduction, etc. After this preliminary analysis there follows an evaluation by synthesis. One should not start directly with synthesis experiments and an incomplete knowledge of the speech wave characteristics. It is helpful to construct alternative hypotheses concerning effective cues already in the analysis stage of the work.

A method of parameter evaluation which has been extensively used in perception research is to make systematic variations of the sound stimulus and determine the

boundary where the response shifts from one phoneme to another. When this technique is applied to several minimally contrasting pairs of phonemes the data can be interpreted on a distinctive feature basis. The absolute values of the boundaries will vary with the particular context of simultaneous, preceding, and following features of the sound matter as well as with prosodic elements. This is the so-called "contextual bias". In an integrated view based on all parameters of importance for a distinction the distinctive feature or rather its speech wave correlate can be conceived of as a vector perpendicular to the hypersurface constituting the multi-dimensional boundary. A similar formulation was given by Chistovich (1967) in her paper at the Congress. The main direction of this vector is the sole remaining attribute of the feature if a common denominator of all possible contexts is to be expressed as was the ambition of Jakobson, Fant, and Halle (1952).

However, a knowledge of this mean direction of the feature vector is not a sufficient end result in speech research. For general descriptive phonetics as well as for automatic speech recognition we need the detailed information of how these boundaries shift with context in the general distributional sense adopted here. The search for formulas enabling us to calculate the contextual bias from the discrete inventory of conditioning factors has already been mentioned in the previous discussion on speech production.

The greater accessibility to the problem from a generative speech production point of view than from a perception point of view has had a certain inhibiting effect on the work at the perception end. We would all agree that the categorization inherent at the production end is quite similar to that at the perception end of the speech communication chain but only defenders of a motor theory of speech perception would argue that perception is nothing but the association of the incoming acoustic stimulus with production categories at the listener's disposal when acting as a speaker. If production and perception categories were identical there no longer remains any difference between a sensory theory and a motor theory but merely a concept of economy in our storage of phonetic categories in a place of the brain common for production and perception.

However, by introspection we can certainly study our own stored sound images of distinctive features and phonemes some of which we might not be able to produce correctly if they belong to a language we are not so well acquainted with. When we mimic speech rapidly the motor activity must be the automatic consequence of a phonetic identification in a previous stage. An identification through what is going on in the efferent motor pathways appears to be an unnecessary complication.

Speech perception is a process of successive and simultaneous identifications in a chain of successively higher levels of language structure. We cannot expect to find a specific brain center for each linguistic category: feature, phoneme, syllable, morpheme, word sentence, but at least a lower level $C(3)$ and a higher level $C(4)$ as proposed in the introductory section. To the inventory belong short term and

long term memory functions as well as feedback mechanism which allow storage comparison and correction. Also it allows a generative prediction of what the speaker is going to say at least at the levels of syntax and semantics, $C(4)$ and $C(5)$. At a level corresponding to a complexity of the order of the syllable $C(2)$ I conceive of an analysis through a window of the width of a few phonemes through which the speech signal passes. I do not hypothesize a strict principle of all phonemes being first identified by their features. Some phonemes are probably identified directly and independent of context, e.g. the (s) . Also the identification is probably not strictly sequential but of arbitrary order within the time span of the window.

This principle overcomes the difficulty of some of the features being specifically sensitive to context. Each identification is a decision based on the probabilities existing at the particular instance and each completed decision influences the distribution of probabilities for the previous and following elements within the window. The general sequential constraints imposed by the language structure and of the speech production mechanism effectively limits the number of alternatives in any decision. This model is also the best principle we can follow in attempts of automatic speech recognition.

The principle indicated above is close to the model of perception outlined by Chistovich in her paper and has an interesting parallel in her experiments on psychological scaling of perceived distances between each of two alternative phonemes and a synthetic sound, the composition of which is varied to produce variations around a perceptual boundary. According to the experiments of Chistovich (1967), Chistovich, Fant, and de Serpa-Leitão (1966) there is some evidence of a gross quantization and scaling at a stage preceding the phonetic identification but this effect is not well established yet.

Vowels, glides, nasals, and laterals appear to offer greater descriptive problems than stops, affricates, and fricatives. One reason is the greater dynamic variability and affinity to coarticulation. The other lies in the variation of scale factors with different speakers. The first and second formant frequencies, F_1 and F_2 , are known to be more important than other parameters, but they are not sufficiently descriptive. The third and higher formants are also of considerable importance in front vowels and serve to differentiate [i] [y] [ɥ] [e] and also [ɛ] and [ø]. F_2 and higher formants of front vowels appear to constitute a single perceptual cue which plays a role similar to that of an F_2 alone in mid and back vowels. This cue is probably not sufficiently specified by a center of gravity only, Fujimura (1967). Spectral width and relative intensity may also be of some importance.

NORMALIZATION OF ACOUSTIC DATA

The average female voice shows 20 per cent higher formant locations than an average male voice and the same average difference is also found between the spectrum patterns of the voices of children (age 8) and female voices, Peterson and Barney (1952). However, the scale factors vary not only with the speaker but also with the

specific vowel and the formant under observation. Thus F_1 of [o] varies but little with the sex of the speaker whereas the scale factor for F_1 of open vowels such as [a] and [æ] are appreciably greater than the average. The origin of these nonuniform variations lies in the non-uniform scaling of the female vocal tract with respect to the male vocal tract, Fant (1966B).

Even if we include all formants in a specification we might find ambiguities such that a female [ø] might have almost the same formant frequencies F_1 , F_2 , and F_3 as a male [e]. Such ambiguities have not been studied in detail but they might be resolved in part by reference to the center of gravity, intensity, and width of the upper group of formants, to a small part by relative levels of peaks and valleys, and in part by a reference to the voice fundamental frequency. It is not known to what extent the normalization with respect to F_0 is a psychological effect, i.e. F_0 acting as a label for the specific female vowel category or whether F_0 enters already in a weighting of the effective timbre. In connected speech we can also expect a normalization with respect to both the immediate and remote context. A related phenomenon is that time variable formants affect the identification more than constant frequency formants. However, it should be appreciated that because of the general relations between formant frequencies on one hand and formant levels and spectral shape factors on the other hand, Fant (1960), a formant F_4 is not audible in a vowel [u] but has a sensation level equal to that of F_1 and F_3 in the vowel [i].

The boundary shift techniques has been successfully adopted by Fujisaki and Kawashima (1967) for an evaluation of the trading relations between the various vowel parameters. We have used this technique for studying the effect on the source level in the region of F_2 and F_3 as a factor influencing the F_2 — F_3 -boundary between [ɥ] and [y]. A 20 dB reduction of F_2 and F_3 intensity level shifted the $(F_2, F_3)^{1/2}$ threshold by no more than 50 Hz, i.e. rather little. However, the probability of [ɥ] identifications rose significantly within the main [y] region.

The extent to which we can approximate a vowel specification by F_0 , F_1 , and a few measures related to an effective upper formant region is not yet determined but is one of the hypotheses that we can test with synthetic speech. For this purpose the upper formant should be generated in parallel with F_1 and shaped with a filter of greater width and selectivity than a simple formant circuit.

CONCLUSIONS

The concepts of distinctive features and cues should be kept apart as belonging to the message inventory and the speech signal inventory, respectively. In search of the physical and psychological reality behind the categorical effects in speech production and perception we might find a system of features constituting a natural ensemble of minimal message units. Such an ensemble can only in part be expected to conform with the system of Jakobson, Fant, and Halle (1952) and I expect it to be more redundant. There remain many questions to be studied concerning the relations between speech parameters and members of a feature or phoneme inventory.

ACKNOWLEDGMENTS*

I am indebted to Björn Lindblom and Sven Öhman for many fruitful discussions on form and contents of this paper.

REFERENCES

- (1966) Borovičková, B. and Maláč, V.: "Towards the Basic Units of Speech from the Perception Point of View", *Proc. of the Seminar on Speech Production and Perception, Leningrad 1966*, 83—88 (Z.f. Phonetik, Sprachwissenschaft und Kommunikationsforschung, 21, Heft 1/2, 1968).
- (1967) Chistovich, L.: "Method of Studying the Decision Rules Applied in Speech Perception", to be publ. in *Proc. of the 6th International Congress of Phonetic Sciences, Prague 1967*
- (1966) Chistovich, L., Fant, G., and de Serpa-Leitão A.: "Mimicking and Perception of Synthetic Vowels. Part II", *STL-QPSR*, No. 3, 1—3 (Stockholm 1966).
- (1960) Fant, G.: *Acoustic Theory of Speech Production* (The Hague 1960).
- (1962A) Fant, G.: "Descriptive Analysis of the Acoustic Aspects of Speech", *Logos*, 5, 3—17 (1962).
- (1962B) Fant, G.: "Sound Spectrography", *Proc. IVth Int. Congr. of Phonetic Sciences, Helsinki*, 14—33 (The Hague 1962).
- (1966A) Fant, G.: "The Nature of Distinctive Features", *STL-QPSR*, No. 4, 1—14 (Stockholm 1966).
- (1966B) Fant, G.: "A Note on Vocal Tract Size Factors and Non Uniform F-Pattern Scalings", *STL-QPSR*, No. 4, 22—30 (Stockholm 1966).
- (1968) Fant, G.: "Analysis and Synthesis of Speech Processes", a chapter in *Manual of Phonetics*, ed. B. Malmberg, (North-Holland Publ. Co., Amsterdam 1968).
- (1961) Fant, G. and Lindblom, B.: "Studies of Minimal Speech Sound Units", *STL-QPSR*, No. 2, 1—11 (Stockholm 1961).
- (1966) Fant, G., Lindblom, B., and de Serpa-Leitão, A.: "Consonant Confusions in English and Swedish—A Pilot Study", *STL-QPSR*, No. 4, 31—34 (Stockholm 1966).
- (1967) Fujimura, O.: "The Spectral Shape in the F2-F3 Region", *Models for the Perception of Speech and Visual Form*, ed. by W. Wathen-Dunn, 251-256 (Cambridge, Mass. 1967).
- (1967) Fujisaki, H. and Kawashima, T.: "Roles of Pitch and Higher Formants in Perception of Vowels", *Digest of the 7th International Conference on Medical and Biological Engineering* Aug. 14—19, 1967, Stockholm, Session 24-2.
- (1952) Jakobson, R., Fant, G., and Halle, M.: "Preliminaries to Speech Analysis. The Distinctive Features and Their Correlates", *Acoust. Lab., M.I.T. Techn. Rep.*, No. 13 (Cambridge, Mass. 1952); 4th printing publ. by The M.I.T. Press (Cambridge, Mass. 1963).
- (1957) Liberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C.: "The Discrimination of Speech Sounds Within and Across Phoneme Boundaries", *J. of Exp. Psychol.*, 54, 358—368 (1957).
- (1963) Liberman, A. M., Cooper, F. S., Harris, K. S., and MacNeilage, P. F.: "A Motor Theory of Speech Perception", Paper D3 in *Proc. of the Speech Communication Seminar, Stockholm 1962*, Vol. II (Stockholm 1963).
- (1967) Liberman, A. M., Cooper, F. S., Harris, K. S., MacNeilage, P. F., and Studdert-Kennedy.

*) This study has been supported in part by Swedish and US governmental funds, more recently by NIH Research Grants NB 04003—05 and HD 02111-02 and by US Air Force Grant AF EOAR 67-34.

- M.: "Some Observations on a Model for Speech Perception", *Models for the Perception of Speech and Visual Form*, ed. by W. Wathen-Dunn, 68—87 (Cambridge, Mass. 1967).
- (1963) Lindblom, B.: "Spectrographic Study of Vowel Reduction", *J. Acoust. Soc. Am.*, 35, 1773—1781 (1963).
- (1966) Öhman, S. E. G.: "Coarticulation in VCV Utterances: Spectrographic Measurements", *J. Acoust. Soc. Am.*, 39, 151—168 (1966).
- (1967) Öhman S. E. G.: "Numerical Model of Coarticulation", *J. Acoust. Soc. Am.*, 41, 310—320 (1967).
- (1965) Öhman, S. E. G. and Lindqvist, J.: "Analysis-by-Synthesis of Prosodic Pitch Contours", *STL-QPSR*, No. 4, 1—6 (1965).
- (1952) Peterson, G. E. and Barney, H. L.: "Control Methods Used in a Study of the Vowels", *J. Acoust. Soc. Am.*, 24, 175—184 (1952).
- (1966) Stevens, K. N.: "On the Relations Between Speech Movements and Speech Perception", *Proc. of the Seminar on Speech Production and Perception, Leningrad 1966*, 102—106 (Z. f. Phonetik, Sprachwissenschaft und Kommunikationsforschung, 21, Heft 1/2, 1968).
- (1966) Stevens, K. N. and House, A. S.: "Speech Perception", a chapter prepared for *Foundations of Modern Auditory Theory*.

DISCUSSION

Akhmanova:

The division into "engineers" and "linguists" which Prof. Fant appears to have laid considerable stress on is the one point where I could not quite follow him—otherwise I could not agree more with Prof. Delattre's comment: this report of Prof. Fant is an excellent piece of work, an invaluable contribution to the progress of phonetic science. I am convinced that the successful development of phonetics depends on an even more minute analysis of what is *actually* going on in linguistic communication. We must emancipate ourselves as much as possible from all kind of phonological (or graphological?) preconception of globality.

Carnochan:

The linguist might profitably move closer to the engineer in accepting the sort of physical features of speech to be evaluated, and make use of different notations for different purposes. A phonemic transcription is always necessary for reading, but for analytical statement, a prosodic approach in phonology may bring out relevant systemic contrasts of the syllable, word and of the longer piece, relating the spectrograms and the phonology in a rather more illuminating way.

Lehiste:

In your lecture you presented ample evidence for the variability of acoustic features and for their interdependence when they function as perceptual cues. Two of the basic premises of the theory of distinctive features were orthogonality and invariance: the distinctive features were to be independent of each other, and the physical manifestations of a given feature were supposed to contain some invariant elements. After your lecture, neither of the premises appears to hold any more.

Pike:

Papyrologists find it necessary sometimes to read a "Whole document" before being able to "read" separate words. Cursive writing, even in English, may smear at the end of words into a scrawl with indistinguishable final letters. Isn't there an analogy here to the acoustic problem?

It seems to me that the experiments here are studying the contrasts at points of maximum differentiation. In normal speech something more like the smearing of cursive writing may also be found.

Zimnyaya:

Perception can be regarded as a working process of at least two channels, which are parallel. In real life we perceive speech on the basis of probabilistic guessing (prognosticating) the whole structure of a word. This guessing may be based on some features, so to say, critical points which evidently are located in the first segment but they may be placed elsewhere too. So the problem of a unit of perception may be regarded as a problem of a part and a whole.

Fant:

ad Akhmanova: I agree with Prof. Akhmanova that we should not build up separate descriptive theories on speech for engineers and linguists. We have a common interest in being able to introduce more or less redundancy in existing specificational systems and to develop new systems that hopefully would conform with an overall generative system.

ad Lehiste: Distinctive features are by definition independent and display combinatory contrasts only. The production-speechwave and perception correlates are in general not orthogonal. Thus in vowels F_1 and F_2 we have to define not only compactness and gravity but also flatness —($F_1 + F_2$). The lack of orthoquality is not important for the theory.

MARKEDNESS*

MORRIS HALLE**

The questions that concern me in this paper are the direct outgrowth of the study of English phonology that N. Chomsky and I have been engaged in for about a decade.¹ This investigation has brought out certain inadequacies in the theoretical framework, and I shall discuss here a proposal that we have made in order to remedy these inadequacies. It is worth remarking that the inadequacies in our theoretical framework were discovered only because we were trying to push the theoretical machinery to its limits, to subject it to the linguist's equivalent of the engineer's break-down test. Such a test, however, can only be conducted by a detailed examination of the functioning of the theoretical apparatus. This in turn pre-supposes that, on the one hand, there be a high degree of clarity as to how a given body of data is to be characterized formally by the descriptive machinery and that, on the other hand, these characterizations, i.e., the phonological rules and their formal properties, be taken fully as seriously as the linguistic facts that they characterize. In the following discussion the phonological rules, therefore, occupy the central position. They are not just more or less accidental formulations, of interest to the linguist only by virtue of the facts that they reflect. They are, rather, the essential organon, the main tool for acquiring deeper knowledge about the phenomena in question.

It is all but self-evident that in every science there are more and less effective ways of describing particular observations. Thus, all of us would normally prefer a simple formula to a list even if the formula and the list cover precisely the same body of facts. Karl Verner is justly esteemed as one of the leading linguists not because he discovered any new facts, any new Germanic forms that were exceptions to Grimm's Law, but rather because he discovered the simple formula —Verner's Law—that succinctly characterizes the list of exceptions. Since different formulations of a given body of data are thus clearly not all regarded as equivalent, it is essential to attempt to discover the general principles on the basis of which one formulation

* This work was supported in part by the U.S. Air Force [ESD Contract AF 19 (628)—2487] and the National Institutes of Health (Grant MH-13390-01).

** Massachusetts Institute of Technology.

¹ The present study summarizes the main points of the final chapter of Chomsky and Halle (1968).

is to be valued over another. As is well known, we have proposed that an appropriate evaluation measure for phonological rules as well as for lexical representations is the number of features required to state them: the lower this number, the greater the generality and the higher the value of the description.² We have supported this proposal by exhibiting certain cases where indeed it led to the correct results. As a very simple example we may cite the fact that (1a) is a more general, and hence more highly valued rule than (1b).

- (1) (a) $[i, e, \text{æ}] \rightarrow [\text{ü}, \text{ö}, \text{œ}]$
 $[-\text{back}] \rightarrow [+ \text{round}]$
- b) $[i, e] \rightarrow [\text{ü}, \text{ö}]$
 $\begin{bmatrix} -\text{back} \\ -\text{low} \end{bmatrix} \rightarrow [+ \text{round}]$

There are, however, quite a number of instances where simple-minded feature counting fails to express the state of affairs properly. As examples consider (2):

- (2) (i) (a) $u \rightarrow i$ (b) $u \rightarrow \text{ü}$
 $\begin{bmatrix} +\text{high} \\ +\text{back} \\ +\text{round} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{back} \\ -\text{round} \end{bmatrix}$ $\begin{bmatrix} +\text{high} \\ +\text{back} \\ +\text{round} \end{bmatrix} \rightarrow [-\text{back}]$
- (ii) (a) $k \rightarrow \text{č}$ (b) $k \rightarrow p$
- (iii) (a) $k \rightarrow \text{č}$ (b) $\text{č} \rightarrow k$

In (2i) the left rule requires more features to state than the right hand rule but it is by no means obvious that there is a real distinction in the generality of the two rules. In (2ii) $k \rightarrow \text{č}$ requires more features than $k \rightarrow p$, but $k \rightarrow \text{č}$ is certainly a more general rule than $k \rightarrow p$. Finally the rule $k \rightarrow \text{č}$ is much more common than the rule $\text{č} \rightarrow k$ but this fact finds no reflex in the number of features that are required to formulate these two rules. Thus, we have here a number of instances where feature counting clearly will not make the sort of choices that an evaluation measure would be expected to perform.

Feature counting fails yet in another way. We have proposed (Halle [1959] and elsewhere) that the regularities in the phonological composition of morphemes in a language would be properly expressed if lexical items were represented by the minimum number of features that allowed the item to be unambiguously identified in the lexicon. The omitted features would then be specified by general rules (morpheme structure rules). As a consequence, in the lexical representations of a language that had only the three vowels /a i u/ these vowels would be specified by fewer features than in a language that had the five vowels /a i u e o/, as shown in (3).

² For a detailed discussion of the problems of the evaluation procedure see Stanley (1967).

(3) (a)

	a	i	u
high	-	+	+
back		-	+

(b)

	a	i	u	e	o
high	-	+	+	-	-
back		-	+	-	+
low	+			-	-

This is, of course, the desired result since in the language (3a), which has fewer vowels, more features are specified by general rule than in (3b), which has more vowels. In the three-vowel language (3a) the feature *low* is determined by the rule that high vowels are nonlow, and nonhigh vowels are low; the feature *low* cannot be determined by rule in the five-vowel language (3b) and hence cannot be omitted in the lexical representations there without leading to confusion among distinct items.

A parallel argument holds for the representations of the vowels in (4),

(4) (a)

	æ	ü	i
high	-	+	+
back		-	+

(b)

	æ	ü	i	ō	γ
high	-	+	+	-	-
back		-	+	-	+
low	+			-	-

with the result that the two vowel systems in (4) are represented by the same set of features as the vowel systems in (3). The only difference between the vowel systems in (3) and (4) is the respective rules specifying the features that have been omitted in the lexical representation. This minor formal difference, however, reflects inadequately the very profound differences that must exist among languages with vowel systems such as those in (3) and those with vowel systems such as (4).

The examples of the failure of the proposed evaluation criterion raise seriously the question whether it is correct to weight all features equally. There appear to be clear cases where some features are more equal than others. Or less facetiously, there are grounds for questioning the completely formal, abstract view of features that has been fundamental to our approach, and that led us, as a matter of principle almost, to neglect the intrinsic content of the features and the rules. The question that posed itself, therefore, was how to take account of this intrinsic content of features and rules.

The direction towards a solution became clearer to us as we attempted to resolve another difficulty that stemmed from our decision to omit features in lexical representations. Lightner (1963) showed that when features were omitted in lexical representations, it was possible to effect certain spurious economies by appropriate

convention (5vb) replaces [u round] by [—round] yielding the feature complex

$$\begin{bmatrix} +\text{low} \\ -\text{high} \\ -\text{back} \\ -\text{round} \end{bmatrix}$$

or [æ]. In (6a) we have given the feature markings of the main vowel types, and in (6b), their phonological feature composition which will be derived from the markings in (6a) in the fashion just described.

(6) (a)

	a	i	u	æ	e	o	ɔ	ü	ĩ	ö	ə	æ
low	u	u	u	m	u	u	m	u	u	u	u	m
high	u	u	u	u	m	m	u	u	u	m	m	u
back	u	—	+	m	—	+	u	—	+	—	+	m
round	u	u	u	u	u	u	m	m	m	m	m	m

(b)

	a	i	u	æ	e	o	ɔ	ü	ĩ	ö	ə	æ
low	+	—	—	+	—	—	+	—	—	—	—	+
high	—	+	+	—	—	—	—	+	+	—	—	—
back	+	—	+	—	—	+	+	—	+	—	+	—
round	—	—	+	—	—	+	+	+	—	+	—	+

We can now return to the problems that arose as a result of our decision to omit predictable features in the lexical representations of morphemes. We recall that one of our reasons for omitting these features was that we wished to extend our evaluation criterion, which consists in counting features in rules, to the lexical representations. We now have another way of evaluating lexical representations. We assume that instead of being represented by matrices such as (6b) in which + and — figure, lexical items are represented by matrices such as (6a) in which features are marked or unmarked. We now postulate that in assessing the cost of a given lexical representation we count the number of feature entries that are not unmarked. Thus, for example, given the matrix (6a) the vowel [æ] has a cost of two features, the vowel [ö] a cost of three and the vowel [a], a cost of zero.

It is to be noted that when segments are represented with feature markings there is no natural way in which one may omit to specify a feature. Thus, in (6a) there are only marked or unmarked, but no omitted features. This fact has bearing on the

solution of the difficulties discovered by Lightner. If lexical representations are framed in terms of marked and unmarked features, we need not have any unspecified features in the lexicon. We recall that the primary motivation for omitting features in the lexicon was that this served as an evaluation measure. We have now an evaluation measure that does not require us to have features unspecified in the lexical representations. We can, therefore, dispense with the procedure of leaving features unspecified in the lexicon, thereby avoiding the serious problems to which Lightner first drew attention.

It may finally be noted that there is an alternative way of looking at the tables in (6). Rather than suppose that the lexical representations are framed in terms of feature markings to be translated by the conventions (5) into plusses and minuses of phonological features, we may suppose that the lexical representations are framed in terms of plusses and minuses and that the function of the conventions (5) is to provide us with a means of assessing the cost of a given configuration of plusses and minuses. We do this by discovering the particular configuration of feature markings that is translated into the appropriate plusses and minuses by the conventions (5) and then counting the number of marked features in the former. It is important that there is no substantive difference between these two procedures; they are notational variants of one and the same thing.

We must now return to the problems that we raised in connection with the vowel systems in (3) and (4). Why is the set [a i u] preferable to the set [æ i ü]? The answer is obvious if we examine the table (6a). The former set has 2 marked features in toto; the latter has 6 marked features. Thus in order to make the proper choice in these cases, we must require formally that in choosing a vowel system, the one least complex in terms of markedness be chosen. Observe that if this is made a formal requirement, then in the lexical representations of a language that has a 3-vowel system, the vowels will always be [a i u]. If the phonological rules require for their input the system [æ i ü] this will be achieved by a redundancy rule that will operate on the lexicon. The difference, therefore, between a language with the system [a i u] and one with [æ i ü] will be that a redundancy rule will be required only in the latter case, but not in the former.

It is, however, readily apparent that complexity as defined here will not provide an answer in all cases. Compare the 5-vowel systems in (7).

(7)

i	u	i u i ü	i u	i u	i ü u
e	o			e	e
a		a	æ a ɔ	a ɔ	a

It is clear that the first is to be preferred over the rest. To formally implement this choice we need two further principles.

(8) Other things being equal, a system in which more features have only the entry u is preferable to a system in which fewer features have only the entry u.

This principle immediately eliminates all but the first two systems in (7). The meaning of (8) is that it prefers the system that makes maximal utilization of each feature for purposes of marking. It, therefore, builds into the theory a bias in favor of *more symmetry*. Observe that the symmetry is of a very abstract sort.

The choice between the first two systems in (7) needs yet another principle. We believe that this principle is that of feature hierarchy. Since we have been unable to discover a complete feature hierarchy, we shall give only a small part of it here as (9).

(9) No vowel system can have a segment marked for round, unless it also has a segment marked for high.

The question of feature hierarchy is certainly of the greatest importance but it has so far resisted all of our attempts to find a proper solution. We see the difficulty immediately when we consider 4-vowel systems. Given the principles enunciated so far, a 4-vowel system must include [*i u a*] plus a vowel of complexity 2 that is not marked for rounding. We see in (6a) that there are 3 candidates that satisfy these conditions [*æ e o*]. It would seem that a proper extension of the feature hierarchy should provide the correct choice but so far we have been unable to discover how to do that.

In spite of the present unsatisfactory situation I think that it is reasonably clear that marking conventions like those in (5) supplemented by certain further principles will enable us to overcome the inadequacy noted in connection with the existence of different vowel systems. It should be underlined that by means of the marking conventions we are taking specific account of the content of the features, for each feature is interpreted by a particular convention in some particular context, in its own specific way.

We must now turn to the second set of examples that we discussed at the beginning, namely the examples illustrating a certain inadequacy in the formulation of rules, given the present theoretical framework. Consider first the examples in (2i). The reason that (2ia) is more complex than (2ib) is that we have no way in the present theory to cause one feature to change in tandem with a change in some other feature. It seems in many instances, however, e.g., in (2i), that this is precisely what happens, when backness changes in many cases it seems normal for rounding also to change. Our problem, therefore, is to discover some device that would allow for such changes in tandem.

It may not be obvious that such a device is already at hand in the marking conventions. In particular, convention (5va) says that in nonlow vowels rounding must agree with backness. Our problem, therefore, is now to bring to bear the marking conventions on the interpretation of individual phonological rules. This can be done in the following way. We can require that whenever a phonological rule changes a feature in such a way that the segment undergoing the change would satisfy the environment of any marking convention, then the segment will also be provided with the feature specification that the convention provides for the unmarked feature.

Thus, a rule

$$[+high] \rightarrow [-back]$$

will automatically also imply $[-round]$ by virtue of convention (5va). However a rule such as

$$[+low] \rightarrow [-back]$$

will not link to any convention since convention (5va) applies only to $[-low]$, and not to $[+low]$ vowels. If the proposed procedure is adopted in general, it will now be rather more difficult to state (2ib), for we shall have to have some way of undoing the effects of the convention (5va). The appropriate procedure would be to add a second line on the right hand side of the arrow so that the rule would read:

$$[+high] \rightarrow \left[\begin{array}{l} -back \\ +round \end{array} \right]$$

The number of features that have to be mentioned in a rule does not in all cases correspond to the number of features that are actually changed by the rule. Thus under the conventions just discussed the rule

$$[+high] \rightarrow [-back]$$

changes both rounding and backness; whereas the more complicated rule

$$[+high] \rightarrow \left[\begin{array}{l} -back \\ +round \end{array} \right]$$

changes only backness and leaves rounding intact.

An even more complicated example is that involved in the Great English Vowel Shift. The Vowel Shift accounts for such alternations as

divine	divinity	profound	profundity
keep	kept	lose	lost
profane	profanity	verbose	verbosity
<i>i</i> ~ <i>ī</i> → <i>āy</i>		<i>ɔ</i> ← <i>u</i> ~ <i>ū</i> → <i>āw</i>	
<i>e</i> ~ <i>ē</i> → <i>īy</i>		<i>ɔ</i> ← <i>o</i> ~ <i>ō</i> → <i>ūw</i>	
<i>æ</i> ~ <i>ǣ</i> → <i>ēy</i>		<i>a</i> ← <i>ɔ</i> ~ <i>ɔ̄</i> → <i>ōw</i>	

In *The Sound Pattern of English* (Chomsky and Halle, 1968) arguments are presented showing that these alternations are actually alternations between tense and lax vowels and that the diphthongization and changes in vowel quality in the tense vowels are due to the operation of special rules, in particular those in (10).

$$(10) (a) \emptyset \rightarrow \left[\begin{array}{l} \text{Glide} \\ \alpha\text{back} \end{array} \right] / \left[\begin{array}{l} V \\ +\text{tense} \\ \alpha\text{back} \end{array} \right] \text{---}$$

$$(b) (i) \left. \begin{array}{l} [\gamma\text{high}] \\ [-\text{low}] \end{array} \right\} \rightarrow [-\gamma\text{high}] \quad / \quad \left[\begin{array}{l} +\text{tense} \\ \alpha\text{back} \\ \alpha\text{round} \\ V \end{array} \right]$$

$$(ii) \left. \begin{array}{l} [\beta\text{low}] \\ [-\text{high}] \end{array} \right\} \rightarrow [-\beta\text{low}] \quad / \quad \left[\begin{array}{l} +\text{tense} \\ \alpha\text{back} \\ \alpha\text{round} \\ V \end{array} \right]$$

The vowel shift rule (10b) abbreviates 2 pairs of disjunctive rules as shown in (11).

- (11) (ai) $\begin{bmatrix} +\text{high} \\ -\text{low} \end{bmatrix} \rightarrow [-\text{high}]$ (aia) $\begin{bmatrix} -\text{high} \\ -\text{low} \end{bmatrix} \rightarrow [+high]$
 (bi) $\begin{bmatrix} +\text{low} \\ -\text{high} \end{bmatrix} \rightarrow [-\text{low}]$ (bia) $\begin{bmatrix} -\text{low} \\ -\text{high} \end{bmatrix} \rightarrow [+low]$

The first three of these rules do not link to any of the marking conventions, hence they have the following effects:

$$\begin{array}{lll} \bar{i}y \rightarrow \bar{e}y & \bar{e}y \rightarrow \bar{i}y & \bar{\alpha}y \rightarrow \bar{e}y \\ \bar{u}w \rightarrow \bar{o}w & \bar{o}w \rightarrow \bar{u}w & \bar{\alpha}w \rightarrow \bar{o}w \end{array}$$

Since (11ai) and (11aia) as well as (11bi) and (11bia) are disjunctive, (11aia) does not affect segments that have undergone (11bi). However, segments that have undergone (11ai) or (11aia) are not exempt from (11bi) or (11bia). Hence, (11bia) applies to (\bar{e} , \bar{o}) which are the reflexes of (\bar{i} , \bar{u}) after having undergone (11ai). These now become [+low]. But this change links these vowels to marking conventions (5iv) and (5vb) with the results shown here.

$$\begin{array}{lll} 11bia & 5iv & 5vb \\ \bar{e}y \rightarrow \bar{\alpha}y \rightarrow \bar{a}y \rightarrow \bar{a}y & & \\ 11bia & 5iv & 5vb \\ \bar{o}w \rightarrow \bar{\alpha}w \rightarrow \bar{a}w \rightarrow \bar{a}w & & \end{array}$$

In other words, the marking conventions state that when \bar{e} , \bar{o} become low it is more natural for them to become simultaneously also nonround and back, (i.e., / \bar{a} /) rather than preserve their original rounding and backness, and the facts of the Great English Vowel Shift as well as analogous facts from other Germanic languages seem to support this empirical claim implicit in the marking conventions.

Convention (5va) illustrates a further interesting property of the marking conventions. Consider a language like Finnish in which there is vowel harmony with respect to the feature *back*. That is, in this language a word will contain only one of the two sets of vowels shown in (12a);

$$(12) (a) \frac{i \ e \ u \ o \ a}{i \ e \ \bar{u} \ \bar{o} \ \bar{\alpha}}$$

which can be accounted for by two rules such as (12b):

$$(12) (b) (i) \ V \rightarrow [\alpha\text{back}] / \#C \begin{bmatrix} V \\ \alpha\text{back} \end{bmatrix} X\text{-----}$$

$$(ii) \begin{bmatrix} V \\ -\text{low} \\ -\text{round} \end{bmatrix} \rightarrow [-\text{back}]$$

Rule (12b) is an abbreviation of two rules:

$$\begin{array}{l} V \rightarrow [+back] \\ V \rightarrow [-back] \end{array}$$

and hence affects [u] turning it into [\bar{u}]. The question now arises whether or not convention (5va) should apply here. Clearly the answer is *no*, for if the convention were to apply, we should have to give the rule in quite complex form. Instead of (12) we should have to have

$$(13) \ V \rightarrow \begin{bmatrix} \alpha\text{back} \\ \beta\text{round} \end{bmatrix} / \#C \begin{bmatrix} V \\ \alpha\text{back} \end{bmatrix} X \begin{bmatrix} \beta\text{round} \end{bmatrix}$$

We observe, however, a further fact: the marking convention (5va) does not apply to all cases to which rule (12bi) applies; rule (12bi) applies to low vowels as well as to nonlow vowels. This observation suggests that the way to avoid the difficulty just discussed is by imposing the constraint (14):

- (14) A marking convention applies either to all or to none of the segments affected by a rule.

Since vowel harmony affects all vowels, whereas (5va) affects only non-low vowels, convention (5va) does not link to the vowel harmony rule, which can therefore be given in the simple form of (12bi).

We have found that (14) correctly handles a number of other cases as well: it provides the correct answer in the various types of umlaut rules, and it also accounts for the appearance of (θ) in place of the less marked (s) in cases such as Grimm's Law and the Semitic spirantization. These are extremely strong bits of evidence supporting the proposed constraint.

The proposed marking and linking conventions are, thus, supported in two distinct ways. On the one hand, they are justified by the fact that the predominant types of phonological system exhibit the properties imputed to them by the proposed conventions. The same conventions appear, furthermore, to be justified by the fact that they account for the predominant types of phonological rules incorporated into the grammar of each individual language. We have here two disjoint sets of facts that appear to be accounted for by a single theoretical extension, the introduction of the markedness conventions. That two disjoint sets of facts appear to be accounted for by the same extensions of the theory, i.e., the introduction of the markedness conventions, is hardly an accident. It suggests that further explorations in this direction are likely not to be totally fruitless.

BIBLIOGRAPHY

- Halle, Morris, *The Sound Pattern of Russian* (The Hague, 1959).
 — with Noam Chomsky, *The Sound Pattern of English*, Harper and Row, Inc., New York, 1968.
 Lightner, Theodore, "A Note on the Formulation of Phonological Rules", *Quarterly Progress Report of the Research Laboratory of Electronics, M.I.T., Cambridge, Mass. No. 68*, January 15, 1963.
 Stanley, Richard, "Redundancy Rules in Phonology", *Language*, 43, 393—436 (1967).
 Trubetzkoy, N. S., *Principles de phonologie* (translated by J. Cantineau), Paris, 1949.

Akhmanova:

The misunderstanding that is bound to arise here is due to faulty terminology: what is proposed here is *not* phonology; it is *morphology*. As such it is very useful and, clearly one of the possible approaches to the problem (particularly acceptable to those who are looking for "universals", for the features that are, supposedly, common to all mankind), But it is completely misleading to call it "phonology". The task of phonology is to discover and determine which of the innumerable possibilities of sound production (the sounds the human organs of speech can produce) are used in different languages as "diacritics" of the "feature level"; in other words, investigate those contrasts and oppositions which are perceived and used by the speakers of different languages.

Martens:

Anmerkung zum Diskussionsbeitrag von Herrn Wiesinger: Es ist methodisch nicht glücklich, "Vokal e" den "normalsten" deutschen Vokal zu nennen (auf Grund der statistischen Frequenz). Eine solche Klassifizierung läßt die Gruppierung nach Akzentuierungsdistribution unberücksichtigt. Man sollte nicht alle (ə), die *nur* in *nicht-akzentuierter* Position auftauchen, zusammenwerfen mit [e], das *nur* in *akzentuierter* Position vorkommt.

Martinet:

What advantage do we derive from such a complicated scheme as the one which has just been presented? I would think that the reason why the vocalic pattern with *i—u* as the highest vowels should be more normal or frequent than the ones with *u* and *ü*, is that *i—u* secure maximal differentiation through an opposition of minimal and maximal front cavities, whereas *u* and *ü* require both medium cavities. It seems that we have here a conflict between those who think that all languages are basically identical and that the duty of linguists is to discover this identity, and those who think that this duty lies in determining what makes a language different from all others within the framework of what is recognized as necessary for an object to be called a language.

Wurzel:

I. Was wollen wir mit der Theorie der Markiertheit?

1. Eine Theorie phonologischer Universalien (Aufbau von phonologischen Systemen) oder
2. Möglichst einfache Lexikoneintragungen und Regeln in *einer* bestimmten Sprache. Beides kann sich sehr stark widersprechen (Beispiel: Phonem [e] im Deutschen häufigster und normalster Vokal).

Wenn wir uns für 2. entscheiden, dann gibt es auch notwendigerweise Markierungswechsel in der Sprachgeschichte.

II. Problem der phonologischen „Nichtakzeptierung“ des germ. Umlauts (das nicht finn.-ugr. oder türk. vorkommt) kann durch Markierung vielleicht gut erklärt werden. Verbreitung der Entrundung gibt es in nahezu allen germ. Sprachen. Gerundete *vordere* Vokale entwickeln sich sekundär als [m rund].

Streben nach Beseitigung der Markierung.

ü, ö sind $\begin{bmatrix} u \text{ dunkel} \\ m \text{ rund} \end{bmatrix}$, Entrundung geschieht nicht durch Wiederzusammenfall mit *u, o*, (denn wie gesagt *vordere* Vokale sind im Deutschen offensichtlich „normaler“. Es wird das

markierte Feature verändert: $\begin{Bmatrix} \ddot{u} \\ \ddot{o} \end{Bmatrix} \rightarrow \begin{Bmatrix} i \\ e \end{Bmatrix}$

i, e haben die features $\begin{bmatrix} u \text{ dunkel} \\ u \text{ rund} \end{bmatrix}$

KOMMUNIKATIVE UND EXTRAKOMMUNIKATIVE BETRACHTUNGSWEISEN IN DER PHONETIK

GEROLD UNGEHEUER

Es soll hier keine Theorie über irgendeinen Bereich der phonetischen Erscheinungen vorgelegt werden. Vielmehr möchte ich die Aufmerksamkeit auf einen Tatbestand lenken, der, obgleich von alltäglicher Evidenz, in den Kommunikationswissenschaften ebenso wie in der Phonetik und Linguistik weithin unbeachtet geblieben ist. Ich möchte zeigen, welche Scheinprobleme aus dem Nichtbeachten dieses Tatbestandes erwachsen und welche Konsequenzen in der Phonetik vor allem aus seiner Existenz zu ziehen sind. Abschließend werde ich das Verhältnis der aus diesen Konsequenzen resultierenden Konzeptionen gegen scheinbar gleiche oder ähnliche in Phonetik und Linguistik zu analysieren versuchen.

Der angedeutete Sachverhalt läßt sich in folgenden Beobachtungen fixieren. Jedes menschliche Individuum macht mit seiner Sprache (oder: hat von seiner Sprache) eine doppelte Erfahrung. Ein erster Erfahrungsbereich wird geschaffen durch die sprachlichen Kommunikationsprozesse selbst: die Individuen kommunizieren mit Hilfe von Sprache; sie erleben dabei Sprache unter ganz bestimmten, eben kommunikativen Bedingungen; sie realisieren Kommunikationsakte und setzen dabei gezielt die ihnen zur Verfügung stehenden linguistischen Mittel ein. Der zweite Erfahrungsbereich ist eingebettet in die vielfältige Aktivität von Wahrnehmungen und Handlungen menschlicher Individuen: sie beobachten sprachliche Kommunikationsprozesse zwischen anderen Individuen, ohne selbst darin verwickelt zu sein, und überschauen in der Retention eigene Kommunikationsakte; in extrakommunikativen Artikulationen nehmen sie Regel- und Gesetzmäßigkeiten der Lautung wahr; durch repetierendes Einüben lernen sie sprachliche Paradigmen und in gleicher Weise verschaffen sie sich einen Überblick über den Wortschatz fremder Sprachen.

In dieser Beschreibung, in der es allein auf die Abgrenzung der beiden Erfahrungsbereiche ankommt, wurde der Terminus „Sprache“ undifferenziert und undefiniert verwendet. Zwei mögliche Interpretationen des Terminus erleichtern das Verständnis der nachfolgenden Erläuterungen. Sprache kann einmal verstanden werden als die Gesamtheit konkreter Kommunikationsprozesse bestimmter Art, die in einer menschlichen Gemeinschaft beobachtet werden; zum anderen kann Sprache interpretiert werden als die Gesamtheit der kommunikativen Mittel, die zur Verwirklichung dieser Prozesse notwendig sind (man könnte von der entsprechenden Grammatik der Sprache sprechen, wenn man Lexikon und Lautung einschließt). Im Hinblick

auf diese beiden Interpretationen läßt sich die zuvor gegebene Beschreibung spezifischer menschlicher Verhaltensweisen in folgende Sätze zusammenfassen.

1. Jedes menschliche Individuum erfährt sprachliche Kommunikationsprozesse in zweifacher Weise, indem es entweder:
 - 1.1 Kommunikation vollzieht (d. h. kommuniziert) und damit als Kommunikator, d. h. als Teilnehmer an einem Kommunikationsgeschehen in Erscheinung tritt,
oder aber
 - 1.2 Kommunikation beobachtet, entweder als externer Beobachter fremder Kommunikationsakte oder als Beobachter eigener, die retentional, d. h. in der Erinnerung, in unterschiedlichen Modifikationen der Wahrnehmung zugänglich sind.
2. Jedes menschliche Individuum hat einen zweifachen Umgang mit den Mitteln der sprachlichen Kommunikation:
 - 2.1 im Vollzug der Kommunikationsakte, indem es sie einsetzt, eine sprachliche Verständigung zu erreichen;
 - 2.2 in manipulierender Beobachtung, indem es sie außerhalb kommunikativer Geschehen zur Kenntnis zu nehmen, einzuordnen und zu klassifizieren versucht.

Die beiden hier konstatierten, menschlichen Verhaltensweisen nenne ich *den kommunikativen bzw. den extrakommunikativen Umgang mit der Sprache*. Mit dieser Unterscheidung ist keinerlei negative oder positive Wertung verbunden. Der Tatbestand wird lediglich als evident festgestellt. Im übrigen kann sofort ergänzt werden, daß für das sprachliche Verhalten der Menschen die extrakommunikative Beschäftigung mit den sprachlichen Mitteln ebenso wichtig ist wie ihr kommunikativer Einsatz.

Es muß erwähnt werden, daß die Differenz zwischen dem Handlungsvollzug unter Mitwirkung eines Werkzeugs und der vergegenständlichenden (objektivierenden) Beobachtung (Wahrnehmung) dieses Werkzeugs in der Phänomenologie Heideggers als wichtige Kategorien unter den Titeln „Zuhandenheit“ und „Vorhandenheit“ abgehandelt wird. In nuce ist in der hier thematisierten Unterscheidung die Dichotomie von Praxis und Theorie angelegt.

Die Termini „kommunikativ“ und „extrakommunikativ“ können leicht mißverstanden werden, wenn man nicht folgende Vereinbarung einhält. Zu jedem Kommunikationsprozeß gehört eine spezifische Umgebung, eine Gesamtheit von Erscheinungen, die ihn einhüllt und zu der mannigfaltige Beziehungen bestehen. Diese Umgebung nenne ich die Kommunikationssituation; ihre Wechselwirkung mit dem Kommunikationsprozeß ist Gegenstand einer kommunikationswissenschaftlichen Ökologie. Die Kommunikationssituation selbst besteht in vielen Fällen wieder aus Kommunikationsprozessen. Hieraus ergeben sich nun Schwierigkeiten

in der Verwendung der Begriffe „kommunikativ“ und „extrakommunikativ“; denn sehr häufig muß ein Sachverhalt, der zu einer einhüllenden Kommunikationsumgebung gehört, in bezug auf den entsprechenden Kommunikationsprozeß als extrakommunikativ, für sich selbst genommen aber als kommunikativ deklariert werden. So enthält der extrakommunikative Umgang mit der Sprache sicherlich wieder kommunikative Akte, da aber diese in Relation zu dem ins Auge gefaßten sprachlichen Kommunikationsgeschehen zur Kommunikationssituation gehören, besteht die Verwendung des Terminus „extrakommunikativ“ gemäß seiner Definition zu Recht. Beide Termini sind relative Bezeichnungen; sie werden verwendet relativ zu demjenigen Kommunikationsprozeß, der Gegenstand der Analyse ist.

Bisher wurden Fakten aus dem sprachlichen Leben der menschlichen Individuen registriert. Es muß nun die Frage beantwortet werden, welche Konsequenzen die gefundene Unterscheidung im menschlichen Verhalten in Phonetik und Linguistik, genauer: für deren Methodologie, Hypothesen und Theorienbildung zu ziehen sind.

Zunächst kann man eine bemerkenswerte Asymmetrie in der Relation der beiden Verhaltensweisen zu wissenschaftlichen Arbeiten schlechthin feststellen. Das Beobachten und Beschreiben von empirischen Daten, auch die experimentelle Forschung können als die wissenschaftliche Ausformung extrakommunikativen Umgangs mit den Phänomenen angesehen werden. In den vielen Bereichen der Wissenschaften, wo es im Objektbereich auf die Unterscheidung zwischen kommunikativem und extrakommunikativem Verhalten nicht ankommt (wohl aber häufig auf den zwischen „Vorhandenheit“ und „Zuhandenheit“), braucht dieser Tatsache keine besondere Aufmerksamkeit gewidmet zu werden. In allen kommunikationswissenschaftlichen Analysen jedoch, also auch in phonetischen und linguistischen, taucht das besondere methodologische Problem auf, daß sich der Wissenschaftler selbst bei seinen Forschungen zwar in die Position der extrakommunikativ handelnden und wahrnehmenden Individuen versetzen kann, die kommunikative Verhaltensweise ihn jedoch sofort aus seinem Beobachtungsturm herausführt, ihn in die menschliche Gemeinschaft zwingt, aus der er nicht mehr als dieses besondere Individuum „Wissenschaftler“ herausragt. Wissenschaft basiert eben prinzipiell auf Beobachtung und Reflektion, sprachliche Kommunikation ist keines von beiden, wenn auch im menschlichen Gesamtverhalten diese wohlpräparierten Kategorien ineinander verschmolzen erscheinen. So ist es nicht erstaunlich, daß in einer kommunikationswissenschaftlichen Disziplin bei Mißachtung dieser Sachverhalte nahezu wie von selbst alle Erkenntnisse auf extrakommunikative Verfahren reduziert erscheinen, wie es tatsächlich in Phonetik und Sprachwissenschaft — das wird noch genauer zu zeigen sein — in weiten Bereichen der Fall ist.

Die wichtigste Folgerung aus der gefundenen Differenz ist daher die, daß bei jeder kommunikationswissenschaftlichen Forschung die Ergebnisse aus einer Analyse extrakommunikativen Verhaltens nicht ohne weiteres und nicht ohne zusätzliche Untersuchungen in den Bereich kommunikativen Verhaltens übertragen werden dürfen, und umgekehrt. Die Forderung lautet: für jeden der beiden Objektbereiche

müssen die adäquaten methodischen Verfahren und die Elemente von Hypothesen- und Theorienbildung unabhängig voneinander aufgebaut werden, bevor die bestehenden Zusammenhänge eruiert werden können. Aus der zuvor erwähnten Charakteristik wissenschaftlichen Arbeitens folgt freilich auch, daß kommunikative Verhaltensweisen und die in ihnen realisierten Systeme von linguistischen Mitteln und kommunikativen Prozeßeinheiten nur extrakommunikativ erforscht werden können. Die darin liegende Schwierigkeit war Hindernis genug, die weitaus meisten Verfahren und Konzeptionen, die in Phonetik und Linguistik bis jetzt bekannt geworden sind, auf extrakommunikatives Verhalten zu beziehen oder zu projizieren. (Auch in der Wissenschaft wird wie in anderen Lebensbereichen bewußt oder unbewußt zuerst das Einfachere abgearbeitet; Schwieriges wird häufig tabuiert, indem man es als uninteressant deklariert.) Die beiden auch in der Phonetik notwendigen methodologischen Ausgangspunkte habe ich im Titel dieses Vortrags, nicht ganz glücklich vielleicht, als kommunikative und extrakommunikative Betrachtungsweisen bezeichnet.

Bevor ich daran gehe, die Rolle des Kommunikativen und Extrakommunikativen in der Phonetik an einigen Beispielen aufzuzeigen, möchte ich kurz darauf hinweisen, daß ich unter Phonetik, moderner Phonetik, die Wissenschaft von allen lautungsbezogenen Erscheinungen der sprachlichen Kommunikation verstehe. Dies schließt die Phonologie als rechtverstandene funktionelle Phonetik ebenso mit ein wie z. B. jede Art diachronischer Phänomene in diesem Bereich. Die Unterscheidung Trubetzkoy's zwischen naturwissenschaftlicher Phonetik und geisteswissenschaftlicher (bzw. sozialwissenschaftlicher) Phonologie halte ich für obsolet. So etwas wie einen „Sprachlaut“ nehme ich freilich nicht als den elementaren phonetischen Gegenstand an; er ist ein abgeleiteter wie andere auch.

Wenn ich nun einige theoretische Elemente der klassischen Phonologie von Trubetzkoy bis Harris näher betrachte, so handelt es sich also in dieser Sicht durchaus um ein phonetisches Thema. Diese Elemente sind sämtlich, gleichgültig ob sie auf distributionelle Analysen angelegt sind oder nicht, von extrakommunikativen Manipulationen linguistischer Einheiten abgeleitet. Es erscheint mir dem Ort dieses Kongresses angemessen, diesen Sachverhalt kurz am Beispiel der phonologischen Theorie Trubetzkoy's etwas genauer herauszuarbeiten.

Im Mittelpunkt der klassischen Phonologien steht der Begriff des Phonems, dessen Definition so angelegt ist, daß sie zugleich die Bedingungen einer „discovery procedure“ erfüllt. Die Phonemdefinitionen enthalten immer zwei Bestimmungskomplexe: einen funktionellen, der den phonologischen Grundgedanken berücksichtigt, und einen stoffgebundenen. Die Definitionen und damit auch die phonologische Analyse gehen aus von einem transkribierten Text, von einem linguistischen Corpus also, das nach rein auditiven Gesichtspunkten segmentiert ist.

Der funktionelle Bestimmungskomplex wird bei Trubetzkoy beherrscht von der „bedeutungsunterscheidenden Funktion“: Phoneme sind die kleinsten Lautsegmente eines Wortes, die eine solche Funktion ausüben. Praktisch wird diese Funktion

festgestellt durch „distinktive Oppositionen“ z. B. bei Minimalpaaren oder nach distributionellen Kriterien (man kann nachweisen, daß die „discovery procedures“ Trubetzkoy's bereits distributionell im Sinne des amerikanischen Strukturalismus arbeiten). Der stoffgebundene Bestimmungskomplex operiert mit dem Begriff der phonetischen Ähnlichkeit. Nur diejenigen nicht in distinktiver Opposition vorgefundenen Lautsegmente werden zu einer funktionellen Klasse zusammengefaßt, die phonetisch ähnlich sind. Jede phonologische Analyse einer natürlichen Sprache führt zu einem Phonemsystem, das zusammen mit den Kombinationsregeln, die die Phoneme erfüllen, den strukturellen Status der Lautungen dieser Sprache charakterisieren.

Man hat nun schon seit langem bemerkt, daß bei der Anwendung dieses phonologischen Verfahrens charakteristische Schwierigkeiten auftauchen, die zum Teil mit anderen Voraussetzungen der Theorie zusammenhängen, die hier nicht interessieren. Zuletzt haben Chomsky und seine Mitarbeiter unter Hinweis auf diese Schwierigkeiten eine Phonologie im Sinne des klassischen Strukturalismus überhaupt in Frage gestellt. Auf diese Auseinandersetzung kann ich hier nicht eingehen. Für mein Thema erscheint mir nur ein Moment dieser Kontroverse wichtig, nämlich die Frage nach dem „Existenzmodus“ von so etwas wie Phonemen. Unter Existenz verstehe ich hier die Rolle solcher Einheiten entweder in einem theoretischen Modell als Konstrukte oder als Elemente der Realität in psychologischer oder physikalischer Interpretation. Eingeschlossen ist darin auch die Weise, in der Chomsky das Wort verwendet, wenn er Phonologien der beschriebenen Art als irrelevant ablehnt mit der Bemerkung: „... the existence of an additional level is highly dubious...“¹

Trubetzkoy hatte sich zuletzt zu einer anti-psychologistischen Position durchgerungen. Phoneme definiert er eindeutig als „abstraktive Einheiten“, und ebenso eindeutig unterscheidet er die Phoneme von ihren Realisationen. Welche Rolle spielen dann aber diese Einheiten im sprachlichen Verhalten der menschlichen Individuen?

Zur Beantwortung dieser Frage müssen einige Merkmale der phonologischen Theorie Trubetzkoy's gesammelt werden. Zunächst bemerkt man, daß der Phonembegriff allein aus extrakommunikativen Betrachtungsweisen hervorgeht. Der Kommutationstest wie auch die Feststellung von Distributionen sind extrakommunikative Manipulationen an extrakommunikativ transkribiertem Material. Die Ergebnisse einer phonologischen Analyse können daher nach den vorhergehenden Erörterungen, wenn überhaupt, nur im extrakommunikativen Verhaltensbereich der Mitglieder einer Sprachgemeinschaft interpretiert werden.

Eine befriedigende Interpretation dieser Art wird nun allerdings erschwert durch die Tatsache, daß der stoffgebundene Bestimmungskomplex in der Definition des Phonems das Moment der phonetischen Ähnlichkeit in fragmentarischer Form ein

¹ N. Chomsky: Current issues in linguistic theory, The Hague, 1964, S. 69.

bezieht. Nicht berücksichtigt wird darin der Sachverhalt, daß Lautsegmente in den verschiedenen Manifestationsbereichen der phonetischen Information — nehmen wir der Einfachheit halber nur den organo-genetischen, den akustischen und den auditiven Bereich — in unterschiedlichen Ähnlichkeitsbeziehungen zueinander stehen. Was bei homorganischer Artikulation einander sehr ähnlich ist, kann auditiv, wie sich in Experimenten herausgestellt hat, sehr unähnlich sein. Dies ist der Fall bei den Lauten [ç] und [j] wie auch bei [v] und [f].²

Ähnliches gilt für jedes Paar von Manifestationsbereichen. Die Konsequenz wäre die, daß bei rigoroser Anwendung des Ähnlichkeitsprinzips verschiedene Phonemsysteme in den verschiedenen Manifestationsbereichen nicht auszuschließen sind.

Die Bestimmung des Phonems als abstraktive Einheit scheint jedoch zusammen mit anderen Formulierungen darauf hinzudeuten, daß Trubetzkoy den Phonembegriff durchaus als theoretisches Konstrukt intendierte, zu dem zwar Korrelate der Realität abgeleitet werden können, das selbst aber nur begriffen werden kann im System einer Strukturbeschreibung, deren Abstraktionsstufe allerdings nicht allzu hoch angesetzt werden darf. Ist diese Deutung des in den „Grundzügen der Phonologie“ (nicht aber in den vorausgegangenen Aufsätzen) angelegten theoretischen Kerns richtig, dann werden im Kommutationstest, auch wenn Informanten hinzugezogen werden, nicht primär psychische oder physikalische Einheiten fixiert; die Feststellung von distinktiven Oppositionen ist vielmehr in erster Linie Glied einer Folge von Manipulationen, die bei der Präparation des Ausgangsmaterials durch eine phonetische Transkription beginnt und die endet bei der Konstitution des theoretischen Konstrukts, genannt Phonem. Die Anwendung des funktionellen Kriteriums, sei es als Bedeutungsunterscheidung oder als distributionelle Komplementarität, ist die eines methodischen Instruments. Das Instrument wird konstruiert nach Prinzipien, die von vornherein als linguistisch sinnvolle Axiome akzeptiert werden. Zu ihnen gehören z. B. das de Saussuresche Linearitätsprinzip und der Vorrang der Distinktion vor der Identifikation.

Was kann aus dieser kurzen Analyse für mein Problem gewonnen werden? Evident scheint dies zu sein: die empirische Basis der klassischen Phonologie ist die extrakommunikative Erfahrung der sprachlichen Kommunikationsmittel. Eine wissenschaftliche Präzisierung dieser Erfahrung wird erreicht in der phonetischen Transkription. Der Kommunikationstest, ebenfalls extrakommunikativer als auditiver Vergleich von Wortlautungen, gehört eigentlich schon einer höheren Ebene an, da er ohne zusätzliche Information mit transkribiertem Material arbeiten kann. Wenn also überhaupt psychische oder physikalische Einheiten aus einer phonologischen Analyse resultieren, so können es nur solche sein, die dem extrakommunikativen Erfahrungsbereich angehören. Für eine Extrapolation ihrer Existenz in den

² G. Ungeheuer: Extensional-paradigmatische Bestimmung auditiver Qualitäten phonetischer Signale, Proc. 5th int. Congr. phon. Sci., Münster 1964, S. 556—560.

kommunikativen Erfahrungsbereich fehlt in den theoretischen Voraussetzungen der Phonologie jegliche Grundlage.

Ist eine solche Real-Interpretation der Phoneme aber notwendig? Wären die phonologischen Theorien gründlich ausgebaut worden, sicherlich nicht. Dann nämlich würde man den Status der Konstrukte innerhalb der Theorie zu bewerten wissen; man würde auch ihre realen Korrelate kennen. Die Frage nach dem realen Existenzmodus würde man dann allerdings nicht aufwerfen. Sie wäre falsch gestellt, weil ein Scheinproblem involviert ist.

Nun enthalten so gut wie alle wichtigen Publikationen zur Phonologie an irgendeiner Stelle Aussagen, Hypothesen und Vermutungen, auch scheinbar experimentelle Befunde über derartige Phonemprojektionen in die Realität. Man kann sie vielleicht alle in die Frage zusammenfassen, die Cohen und van Katwijk zu Beginn ihres Beitrags zum letzten Kongreß der Phonetischen Wissenschaften gestellt haben: „Do phonemes have any kind of existence to the extent that they can be perceived?“³

So wie die Frage formuliert ist, gibt es nur eine Antwort: Phoneme als abstraktive Einheiten können nicht perzipiert werden. Wahrgenommen werden höchstens Phonemrealisationen. Aber die Frage soll wohl gar nicht so wörtlich verstanden werden. Gemeint ist vielleicht, ob Funktionsgleichheit und -ungleichheit von den Mitgliedern einer Sprachgemeinschaft gerade so erfahren werden, wie es aus einer phonologischen Analyse hervorgeht. Empirisch läßt sich diese Frage natürlich nur beantworten, wenn man diesen Individuen das Lautungsmaterial in derselben Form vorlegt, wie es für eine phonetische Transkription zum Zwecke einer phonologischen Analyse angemessen wäre, und zwar vorlegt zum Zwecke einer extrakommunikativen Beobachtung. Die Antworten der Informanten werden dann aber in vollem Umfange vom Entwicklungsstand ihrer Erfahrung mit extrakommunikativen Manipulationen an sprachlichem Material abhängen. Man wird Antworten bekommen, die phonologische Ergebnisse bestätigen, man wird aber auch solche erhalten, die phonologisch nicht interpretierbar sind. Viele Informanten werden, das ist meine Erfahrung, Allophone als verschiedene Einheiten ansehen, andere werden in der Segmentierung gar nicht so weit gehen wollen, wie es die Phonologie verlangt. Man darf auch nicht vergessen, daß eine ausgeprägte Fertigkeit im Segmentieren von sprachlichen Lautungen — wiederum extrakommunikativ — durch das Erlernen und ständige Verwenden einer Buchstabenschrift uns allen in überstarkem Maße zu eigen ist. Wie steht es aber in diesem Punkte mit Individuen, die keine Buchstabenschrift oder überhaupt keine Schrift besitzen?

Wichtiger als dieses Problem ist mir jedoch ein weiteres Moment. In der Frage von Cohen und van Katwijk ist noch die Teilfrage enthalten: welche Rolle spielen Phoneme überhaupt im Kommunikationsprozeß selbst? Dies ist für das vorliegende

³ A. Cohen, A. F. V. van Katwijk: The perception of phonemes as a function of acoustic and distributional cues, Proc. 5th int. Congr. phon. Sci., Münster 1964, S. 231.

Problem eine entscheidende Frage. Hier kommt mit vollem Gewicht die Unterscheidung von kommunikativen und extrakommunikativen Betrachtungsweisen ins Spiel.

Bereits vor dem Auftauchen der Phonologie gab es in der Phonetik den Streit über Existenz oder Nichtexistenz von Sprachlauten, der dann später auf die Phoneme übertragen wurde. Einige Forscher sprachen sich für, andere gegen Laute aus; freilich wird der eigene Standpunkt nicht immer mit aller Schärfe herausgearbeitet, und manches erkennt man erst an den Argumenten, die den Irrtum der anderen erklären sollen. Paul gehört zu denjenigen, welche die Existenz von Sprachlauten in Zweifel gezogen haben. In seinen „Prinzipien“⁴ schreibt er:

„Eine wirkliche Zerlegung des Wortes in seine Elemente ist nicht bloß sehr schwierig, sie ist geradezu unmöglich. Das Wort ist nicht bloß eine Aneinandersetzung einer bestimmten Anzahl selbständiger Laute, von denen jeder durch Zeichen des Alphabetes ausgedrückt werden könnte, sondern es ist im Grunde immer eine kontinuierliche Reihe von unendlich vielen Lauten...“

Und bei Porzig, der durchaus in dieser Tradition steht, liest man.⁵

„Es werden also aus der unendlichen Fülle des tatsächlichen in der Lautung gegebenen einige wenige Züge ausgewählt, auf die es für die Anerkennung eines Lautes als desselben Lautes ankommen soll. Das nennt Karl Bühler das ‚Gesetz der abstraktiven Relevanz‘, wir können auf Deutsch ‚Gesetz der bedeutsamen Auswahl‘ sagen. Für was aber ist diese Auswahl bedeutsam? Nicht für den einzelnen Laut, denn so etwas gibt es gar nicht als selbständige Einheit. Wenn wir bisher von Lauten in diesem Sinne gesprochen haben, so war das eine Anpassung an unser Schulwissen. In Wahrheit sind die Laute der menschlichen Sprache nur wirklich als Züge oder Merkmale an der Lautgestalt des Wortes, so wie Flächen, Kanten und Ecken nur wirklich sind als Züge an einem Würfel, losgelöst von diesem aber gar nicht als für sich bestehend gedacht werden können.“

Zur Zeit des Prager Kreises traten der Prager Phonologie, wie ich die Phonologie Trubetzkoy's und seiner damaligen Kollegen nennen möchte, eine Reihe von Wissenschaftlern entgegen, die mit den dort entwickelten theoretischen Grundlagen sich nicht ganz einverstanden erklären konnten. Zu ihnen gehörte de Groot, Menzerath, Sotavalta und Collinder. Sie stimmten zwar mit den Verfechtern der Prager Phonologie darin überein, daß eine „funktionelle Klangwissenschaft“ (de Groot) geschaffen werden müsse. Der Prädominanz des Distinktionsbegriffs, der die Phonologie Trubetzkoy's charakterisiert, konnten sie jedoch nicht zustimmen. Die Evidenzen nun, die sie für die Rechtfertigung ihrer Kritik beibringen, entstammen alle, wie es auch bei Paul und Porzig der Fall war, dem Bereich kommunikativer Erfahrung mit den Sprachmitteln. Ich zitiere hier nur als Beispiel de Groot:⁶

⁴ H. Paul: Prinzipien der Sprachgeschichte, 1920^s, S. 51/52.

⁵ W. Porzig: Das Wunder der Sprache, München 1960, S. 52/53.

⁶ A. W. de Groot: Phonologie und Phonetik als Funktionswissenschaften, *TCLP* 4 (1931) 116—147, S. 125.

„Die wesentliche Funktion des Phonems ist: dadurch, daß es wiedererkannt, identifiziert wird, nötigenfalls die Wiedererkennung, Identifizierung von Wörtern oder Wortteilen, welche Symbolwert haben, zu ermöglichen oder zu erleichtern.“

Im Grunde ging es also darum, daß der Prager *Distinktionsphonologie* eine *Identifikationsphonologie* entgegengesetzt wurde. Dank der Tatsache, daß sich Phonetik und Linguistik von jeher überwiegend um Befunde extrakommunikativer Evidenz gekümmert hatten, und dank der großen theoretischen Leistung der ersten Phonologen, entwickelte sich die Distinktionsphonologie zur herrschenden Funktionsphonetik. Die Wissenschaftler, die mehr oder weniger explizit für eine Identifikationsphonologie als funktioneller Klangwissenschaft des kommunikativen Sprachumgangs eintraten, blieben in heuristischen Überlegungen stecken. Zum Teil lag dies daran, daß sie ihr Ziel nicht präzise genug herausarbeiten konnten; zum anderen Teil wurden identifikationsphonologische Analysen immer wieder durch den Versuch getrübt, Phoneme und Phonemsysteme der Distinktionsphonologie kommunikativ zu interpretieren.

Bei genauer Lektüre der gegenwärtigen phonetischen und linguistischen Veröffentlichungen ergibt sich, daß die Auseinandersetzung zwischen Distinktions- und Identifikationsphonologie auch heute noch anhält. Ebenso ist es immer noch sehr geläufig, extrakommunikative Ergebnisse in kommunikative umzudeuten, häufig in äußerster Simplizität. Als Beispiel für solche Verfahrensweisen zitiere ich die folgenden, resümierenden Sätze von Halle und Stevens:⁷

„In line with the traditional account of speech production, we shall assume that the speaker has stored in his memory a table of all the phonemes and their actualizations . . .

The analysis procedure that has enjoyed the widest acceptance postulates that the listener first segments the utterance and then identifies the individual segments with particular phonemes.“

Halle und Stevens fassen zwar in diesen Sätzen nur die Auffassung der Mehrheit der Wissenschaftler zusammen, doch enthält ihr Modell am Anfang und Ende des intendierten Kommunikationsprozesses wiederum nichts anderes als Konkatenationen von Phonemen.

Am ausgeprägtesten und bewußt in der Nachfolge de Groot's wird der Standpunkt der Identifikationsphonologie heute von Mol vertreten. Die Reihe seiner zusammen mit Uhlenbeck in *Lingua* veröffentlichten Aufsätze sind wohl vielen in Erinnerung.⁸ Erwähnen muß ich auch den Vortrag, den Mol auf dem letzten Kongreß in Münster unter dem Titel „Are Phonemes really realized?“ gehalten hat. Dort kommt er zu der kategorischen Aussage:

⁷ M. Halle, K. N. Stevens: Speech recognition, a model and a program for research; in: J. A. Fodor, J. J. Katz: The structure of language, 1964, S. 604—612.

⁸ Besonders ist zu nennen: H. Mol, E. M. Uhlenbeck: Hearing and the concept of the phoneme: *Lingua*, 8 (1959) 161—185.

„Phonemes, however, have not been discovered, they have been invented, which makes all the difference.“

Hier erinnert man sich an die Formulierung von Menzerath:⁹

„Ein sozusagen autonom fungierendes Lautsystem gibt es jedoch nicht. Jede Darstellung dieser Art ist offenbar eine Fiktion, ein Spiel mit Abstraktionen ... (S. 123).

Zugrunde liegt bei der Wortbildung nicht der ‚Gegensatz‘, sondern umgekehrt die ‚Formregel‘ ... (S. 125).

In diesen Reden und Gegenreden geht manches durcheinander. Die Situation klärt sich, so glaube ich, nur dann, wenn man den doppelten Umgang der menschlichen Individuen mit den sprachlichen Mitteln als Faktum festhält und diese Differenz prinzipiell auch in die Theorie einführt. Die linguistische Bedeutung der Distinktionsphonologie kann, auch wenn ihre Theorie angreifbar ist, durch kommunikationsbezogene Argumente nicht erschüttert werden. Eine eigenständige, funktionelle Phonetik der kommunizierenden Individuen kann andererseits ihre Grundlage nicht aus dem extrakommunikativen Erfahrungsbereich derselben Individuen gewinnen. Beide Betrachtungsweisen, die kommunikative und die extrakommunikative, sind in der Phonetik gleich wichtig, ihre kategorialen Unterschiede müssen jedoch mit aller Klarheit festgehalten werden. Die Ergebnisse aus den beiden Forschungsbereichen stehen nicht — wer würde dies auch vermuten? — beziehungslos nebeneinander; die Beziehungen können aber erst adäquat analysiert werden, wenn der wissenschaftliche Wert dieser Ergebnisse nicht durch Kontamination kommunikativer und extrakommunikativer Gesichtspunkte herabgesetzt werden.

Wahrscheinlich war Bühler derjenige, der das vorgelegte Problem unter allen seinen Zeitgenossen am besten durchschaut hat. Er war auch derjenige, der kommunikativen Vollzug und extrakommunikative Beobachtung wohl unterscheidend, eine Überbrückung beider Aspekte versuchte. Leider sind die Gedanken Bühlers heute fast in Vergessenheit geraten. Aus diesem Grunde erlaube ich mir, Bühler in einem längeren Zitat zu Wort kommen zu lassen:¹⁰

„Nur muß, wenn auf dem Wege des Signalementverfahrens der Horizont eines mühelosen Wiedererkennens auf Tausende von Einheiten erweitert werden soll, eine zweite Bedingung erfüllt sein, die man nicht aus dem Auge verlieren darf; man kann sie als Psychologie der aufstrebenden Phonologie unserer Tage nicht oft genug ins Merkbuch schreiben. Es ist die einfache Tatsache, daß kein Mensch imstande ist, Tausende von Gebilden, die wie die Eier in unserem Exempel einzig durch Notae-Kombinationen charakterisiert wären, praktisch so spielend, schnell und sicher auseinanderzuhalten, wie das jeder normal geübte Partner einer Sprachgemeinschaft mit den Klanggebilden der Wörter fertig bringt. Das ist eine Behauptung, welche ich zwar

nicht experimentell bewiesen habe, aber aus der Analyse des Wiedererkennens beim Lesen und vielen anderen Daten ableite; ein Faktum, das wie andere erkannt und respektiert sein will und auf die weitgehende Mitwirkung des akustischen Gesichts der Klangbilder bei ihrer Diakrise hinweist. Die Phonologie von heute löst die Aufgabe einer systematisch aufgebauten Diakrisenlehre nur im ersten Schritt und wird beim zweiten zur Gestaltpsychologie in die Lehre gehen müssen.“

Abschließend möchte ich mich nun noch der Frage zuwenden, ob nicht etwa die Unterscheidung von kommunikativer und extrakommunikativer Betrachtungsweise ganz oder teilweise enthalten ist entweder in der de Saussureschen Unterscheidung von „parole“ und „langue“ oder in der Differenz von „competence“ und „performance“. Die Frage wird besonders durch den Sachverhalt herausgefordert, daß sowohl „parole“ wie auch „performance“ sich eindeutig auf die konkreten Akte sprachlicher Kommunikation beziehen, also offensichtlich kommunikativen Konzeptionen zugehören.

Wenden wir uns zuerst der de Saussureschen Dichotomie zu. Es wurde von mehreren Autoren mit Recht darauf hingewiesen, daß der Begriff der „langue“ im „Cours de linguistique générale“ (der ja nicht von de Saussure selbst veröffentlicht, sondern von Bally und Sechehaye herausgegeben wurde) nicht eindeutig beschrieben ist und daher unterschiedliche Interpretationen zuläßt. Die verbreitetste Interpretation ist wohl die, welche den Gesichtspunkt des Systems („System von Systemen“) betont, und die im System gebundenen Einheiten durch negative Abgrenzungen gegeneinander („valeurs linguistiques“) abhebt. In diesem Sinne hat Trubetzkoy „langue“ als „Sprachgebilde“ übersetzt. Man kann aber im „Cours“ auch Stellen finden, in denen die Einheiten der „langue“ als psychische Größen aufgefaßt werden, die voll konkretisiert und nicht nur die linguistisch relevanten Unterscheidungsmerkmale mit sich führend sprachliche Kommunikationsprozesse konstituieren. Im Begriff der „langue“ sind also sowohl kommunikative wie extrakommunikative Aspekte berücksichtigt.

Was die „parole“ angeht, so kann man sagen, daß zwar dieser Terminus sich eindeutig auf die kommunikativen Akte bezieht, der dahinter stehende Begriff von de Saussure jedoch in erster Linie dazu verwendet wird, um den für ihn linguistisch wichtigen Gegenstandsbereich der „langue“ abzugrenzen. Der Gegensatz „kommunikativ-extrakommunikativ“ wird nicht thematisiert. Im Vordergrund steht vielmehr der Unterschied von System und Systemrealisierung in konkreten Akten sprachlicher Kommunikation. Das System, die „langue“ wird nach Kriterien, die dem extrakommunikativen Umgang mit der Sprache nachgebildet sind, aufgebaut, und die Sprachprozesse selbst werden erklärt als geordnete Folgen dieser Einheiten, die aus dem „langue“-Inventar abgerufen und je nach individueller Psyche eingetrübt oder angereichert werden mit zusätzlichen Vorstellungen, — ein Modell, das auch heute in der Linguistik noch sehr lebendig ist. Wenn man also auch in der Differenz von „langue“ und „parole“ viele verdeckte Hinweise auf eine Unterscheidung von kommunikativer und extrakommunikativer Spracherfahrung, bzw. ihrer theoretischen Präzisie-

⁹ P. Menzerath: Die Architektonik des deutschen Wortschatzes; *Phonetische Studien* 3, Bonn 1954.

¹⁰ K. Bühler: Sprachtheorie, Jena 1934, S. 282—283.

zung finden kann, so kann doch von einer vollen Übereinstimmung keine Rede sein.

Ähnliches trifft für das Begriffspaar „competence“ und „performance“ zu. Chomsky hat darauf hingewiesen¹¹, daß dieses in Beziehung steht zu der de Saussureschen Dichotomie. Sehr äußerlich ist bereits insofern eine Übereinstimmung festzustellen, als weder bei Chomsky noch bei einem seiner Mitarbeiter hinreichende Information darüber zu finden ist, was mit den beiden neu eingeführten Kategorien eindeutig gemeint ist. Nach Chomsky entspricht die „competence“ der „langue“ bis auf die Eigenheit, daß in der neuen Konzeption der „kreative“ oder eben „generative“ Aspekt in Form von Regelsystemen dominiert. In der Tat wurde in der generativen Grammatik eine Formalisierungsstufe und damit eine Höhe der Präzisierung erreicht, wie man sie zuvor in der Linguistik nicht gekannt hatte. Die sehr reduzierte, manchmal recht vage, immer aber außerordentlich stereotype Ausdrucksweise hat in den diesbezüglichen Veröffentlichungen jedoch nicht dazu beigetragen, die Grundkategorien der linguistischen Theorie wie „competence“ und „performance“ in allen Bestimmungstücken begrifflich zu fixieren.

Wie bei de Saussure der Begriff der „langue“ macht hier der Begriff der „competence“ offensichtlich die größten Schwierigkeiten. Er wird stereotyp erläutert durch das Wissen („the speaker-hearer's knowledge of his language“), das die Individuen von ihrer Sprache besitzen. Aber gerade am Begriff der „knowledge“ ist nicht klar, auf welchen Erfahrungsbereich er bezogen ist. Seit die „competence“ von Chomsky auf den „idealisierten Sprecher-Hörer“ bezogen wird, ist meines Erachtens die Verwendung von „knowledge“ in diesem Zusammenhang noch unklarer geworden. Die Redeweise vom „idealisierten Sprecher-Hörer“ soll offensichtlich auf diejenigen Züge an den sprachlichen Erscheinungen verweisen, die in einer Sprachgemeinschaft allgemeine, d. h. soziale Geltung haben, während die „performance“ den individuellen, konkreten Vollzug der Sprachakte betrifft, — ein Moment, das auch bei de Saussure eine bedeutsame Rolle spielte.

Zweideutig ist auch der Begriff der Regel, von dem einerseits gesagt wird, daß er innerhalb der Theorie nicht psychologisiert werden darf („generativ“ nicht gleich „produktiv“), der jedoch andererseits (freilich nicht bei Chomsky selbst) auch in Beschreibungen der Performance auftaucht.

Es bleibt mir kein Raum, an dieser Stelle meine Analyse der generativen Begriffsbildungen zu vertiefen. Andeuten wollte ich nur, daß auch „competence“ und „performance“ nicht den Gegensatz von „kommunikativ“ und „extrakommunikativ“ thematisieren, wenn auch viele Momente darin auf ihn verweisen.

Kurz erwähnen möchte ich noch, daß es eine Gegenüberstellung zweier menschlicher Verhaltensweisen gibt, in welche das von mir analysierte Begriffspaar eingeordnet werden kann. Ich meine die Dyade „image“ und „plan“, die von Miller, Galanter

und Pribram diskutiert wurde.¹² In ihrem Buch, in dem man an vielen Stellen sehr deutlich den Einfluß von Ideen aus dem Bereich der generativen Grammatik feststellen kann, wird diese Dichotomie auf weiten Strecken gekoppelt mit Überlegungen, die völlig mit meiner Unterscheidung von extrakommunikativem und kommunikativem Verhalten übereinstimmen. Da sich die Autoren ein viel allgemeineres Thema als nur die linguistischen Verhaltensweisen vorgenommen haben, sind auch ihre begrifflichen Kategorien allgemeiner. Ich meine jedoch, daß dieser Ansatz, befreit von allzu direkter Einflußnahme florierender linguistischer Theorien weiter verfolgt werden sollte.

Ich möchte am Schluß meiner Darlegungen an meinen ersten Satz erinnern: keinesfalls wollte ich hier eine Theorie über irgendeinen Bereich der Phonetik vorgelegen: ich wollte nur aufmerksam machen auf einen wichtigen Tatbestand, der mir evident erscheint, und dessen Bedeutung für Phonetik und Linguistik nach meiner Einsicht fundamentalen Charakter hat. An vielen Stellen meiner Erörterungen wird der Wunsch nach breiterer Beschreibung der Fakten, nach Präzisierung der wissenschaftstheoretischen (epistemologischen) Grundlagen und nach eingehender Begründung meiner kritischen Bemerkungen wach geworden sein. Ich hätte mir gewünscht, die Zeit dafür gehabt zu haben.

DISCUSSION

Akhmanova:

It is extremely difficult to understand why Prof. Ungeheuer speaks so slightly of what he describes as „Distinktionsphonologie“. The Prague School with its numerous developments, particularly, to my mind, the Soviet phonologists with Prof. Avanesov at their head have done excellent work for various languages with this approach to the problem as the general theoretical basis of their work. Phonology is „functional phonetics“. It is concerned with the „distinctive function“ of speech sounds, the „phoneme“ or rather „phonemes“ being two or more sounds the difference between which is not conditioned by position alone.

What Professor Ungeheuer describes as the possibility for an individual „seine sprachlichen Mittel zu beobachten, sie spielerisch auszuprobieren...“ etc., these facts I would relegate to the field of „paralinguistics“ which has been so profitably elaborated by people like Trager in America Crystal and Quirk in England, Renský in this country.

Lüdtke:

Die relativen Positionen von Frau Achmanova und Herrn Ungeheuer lassen sich vergleichen mit denen in der Physik zwischen Newton und Einstein. Der Frage „what is wrong about distinction phonemics?“ kan man in Parallele stellen „what is wrong about Newton's mechanics?“ — Sowohl die Distinktionsphonologie als auch Newton's Mechanik sind *brauchbar* innerhalb eines bestimmten Rahmens und nur dort. Sie sind nicht allgemein gültig. Ob sie richtig oder falsch sind, ist eine müßige Frage.

¹¹ N. Chomsky: *Aspects of the theory of syntax*; Cambridge (Mass.) 1965.

¹² G. A. Miller, E. Galanter, K. H. Pribram: *Plans and the structure of behavior*; New York 1960.

Martinet:

J'aurais aimé que le conférencier nous indique rapidement quelles seraient les conséquences, en matière de méthodes et de résultats, du remplacement de la phonologie distinctive par une phonétique fonctionnelle sur la base de l'identification.

Vachek:

The "extrakommunikative Betrachtungsweise" contains, in Prof. Ungeheuer's definition, things which do not lie on the same level; "spielerisches Ausprobieren" is a thing profoundly different from observation and classification which are necessary attributes of any scientific approach.—The approach of Menzerath and others, stressing the continuous physiological and acoustic flow of utterances does not take into consideration the obvious fact that speakers and listeners quantize this flow (and the emphasis laid on this fact is the main merit of Prague and other phonologists.)—Finally, de Groot's stress on the "identifying function" of phonemes is not quite appropriate—the identifying function certainly exists but it is just a consequence of the distinctive function.

Zwirner:

Die Unterscheidung von kommunikativem und extrakommunikativem Aspekt spielt, wenn auch mit anderen Worten, eine bedeutende Rolle in der Theorie der Geschichtsforschung, insbesondere in der neutestamentlichen Hermeneutik. Seit die neutestamentlichen Schriften unter streng historischem Gesichtspunkt untersucht werden, der im vorigen Jahrhundert eine Methode gefunden hat, muß unterschieden werden, zwischen einem eschatologischen Verständnis dieser Texte im Sinn ihrer Autoren und einem wissenschaftlichen Verständnis dieser Texte im Sinn der modernen Geschichtsforschung. Diese Texte können in diesem Sinn nur verstanden werden, indem die Historiker sich dem Anspruch dieser Schriften im Hinblick auf ihre historische Aufgabe entziehen und sich also „extrakommunikativ“ verhalten. Dabei ist allerdings zu bedenken, daß es sich in diesen beiden Fällen des kommunikativen und extrakommunikativen Verhaltens um den Sinn jener Schriften handelt, zu dessen Begriff es gehört, daß er zusammengefaßt werden kann, und zwar bis zu einem Satz oder sogar bis zu einem Wort. Etwas Ähnliches gibt es für die Linguistik jedoch nicht.

The Present-Day Tasks of the Phonetic Sciences

(Round Table Discussion — September 8)

PROF. D. B. FRY (London)

Perhaps I may be allowed to start the discussion by saying briefly what appear to me, as an experimentalist, to be some of the present-day tasks of the phonetic sciences. In order to do this I shall need to put before you a very general scheme indicating the various stages in our experimental studies and also to cast a rapid glance at some of the things that have been achieved so far.

We begin in every case with observational data, obtained by making observations of speech operating in a perfectly natural environment, as far as this is possible. These data are the result of such things as frequency analyses of the acoustic wave, the study of wave-forms, studies of the physiology of speech by means of X-ray photography, electromyography and similar methods. In experimental work, no step can be taken until we have such data.

Fortunately over the past years a mass of such information has been collected, at least with regard to the phonological or phonemic level. We have, for example, a good deal of spectrographic information and of X-ray information with reference to quite a range of languages. I do not mean to imply that we have enough—it is indeed doubtful whether we shall ever have enough—but we certainly have a good basis for experimentation where the phonemic level is concerned. It would not be true to say the same, however, with respect to the prosodic features in speech, and here I am suggesting one of our present-day tasks. We do not have enough direct measurements and observations in the area of prosodic features and a good proportion of the data we have are not particularly well organized. To take as an example tone and intonation and their relation to fundamental frequency, we need a more systematic approach to observations in this area and in particular a much sharper awareness of the different functions of affective and grammatical intonation. Let me say, in passing, that one of the great pleasures afforded by this Congress was that of hearing Dr. Louise Kaiser talking to us on this very point yesterday. It is of little use to go around collecting observational data about patterns of fundamental frequency variations without keeping clearly in our minds that we must be prepared for at least two large classes of phenomena, those associated with grammatical intonation and those associated with affective or emotional intonation. Given this and further necessary sub-categories of functional intonation patterns, we then need a considerable body of observations of fundamental frequency variation. For instance, I do not think there is anyone at

present who can make statements based on measurements concerning the range and distribution of fundamental frequency curves associated with any grammatical intonation pattern in English. In this connection I ought to say, however, that Professor Delattre has fairly recently published data on features of fundamental frequency variation which characterise intonation differences between several European languages.

Once we have observational data as a basis for experimental work, the next step is to find out how the features which appear in the data are used by the people who employ the particular language. At this stage we come to the necessity for investigating the operation of single acoustic cues. In his paper at the first session Professor Delattre gave us a very fine exposé of the ground that has been covered in the last fifteen or twenty years in this respect and made it clear that we are reasonably well provided with information about the operation of single cues with respect to phonemic systems.

In speech, however, acoustic cues operate in combination and the next task is to find out how various cues interact with each other. When people are talking to each other a multiplicity of cues is operating all the time and I would put forward as one of our major tasks at present the discovery of how a whole system of cues is integrated and used. On this subject we had the pleasure of hearing from Professor Čistovič an extremely interesting and illuminating paper which did more than point the way in this matter of the interaction and interdependence of cues. Included in this area is of course the question of the relative weight given to any cue in a complex of cues by the language users. This again is an immediate task for us and we are very glad to know that Professor Čistovič and her collaborators are already tackling the problem.

In experiments dealing with single acoustic cues we rely a great deal on the use of synthesised speech because this is the only way in which we can control the speech stimuli closely enough. But having worked through single cues to combinations of two and more cues and then to a whole complex of cues, I believe that we must return in a very specific way to experiments with natural speech. This is a somewhat vexed question and I am of course only expressing a personal opinion here. It seems to me one of our tasks is to become much more sophisticated and to discover effective methods of watching the operation of these cues when listeners are dealing with natural speech. Let me say at once that I do not know exactly how this will be done, but I might perhaps give one example from my own work. A good many years ago I used synthetic speech in experiments on stress judgments where the duration of certain stretches of the speech was varied systematically. Methods of controlling the parameters of acoustic stimuli have become much more sophisticated in recent years, both through the use of digital and of analogue methods and in our own laboratories Dr. Huggins has devised a mechanism which will enable me to repeat those experiments by varying durations in utterances which are otherwise natural speech. I myself believe that such experiments are a necessary stage in our study of the way in which speech works.

One further task, and an extremely difficult one, is discovering how all the processing of acoustic information by the listener is integrated with the linguistic side of speech. We need to know how these two things fit together and also how both are related to the statistical information which is continuously applied in speech processing. It would certainly be extremely helpful if someone could invent a machine which would enable us to read out the statistical information which we all have tucked away in the cortex and which each of you is employing at this moment while listening to me. This may be asking too much, but I am sure that we shall in the future achieve some success in finding out how acoustic, linguistic and statistical information are combined in speech processing.

This covers most of what I wanted to say about the tasks which face the experimentalist but perhaps I may be allowed to add one footnote. There is a very urgent need for work on the scales which can be appropriately applied to the input of speech—that is to say the perceptual scales, and this again is something which Professor Čistovič mentioned in her paper. With respect to the dimension of sound quality, for example, we do not know what is the character of the scales which can properly be used; if the quality of a sound is changed by varying some acoustic parameter, we do not know what would constitute a further change of the same magnitude for a given listener, or what would be double the change of quality and so on. In short the form of these scales is something which has really not been tackled at all and until this is done we shall not be able to solve the problem of assigning a particular weight to an acoustic cue or of comparing weights in a meaningful way.

PROFESSOR B. MALMBERG (Lund):

Mr Chairman, Ladies and Gentlemen,

I was asked only a few hours ago to take part in this round table discussion, so I have not prepared anything particular, nor have I had time enough to think over what I should like to say. Nevertheless, we all have problems in our daily work as scholars and teachers, and since I have this opportunity of taking to the members of the Congress I will try to sum up a few points which seem to me to be essential and which should be discussed in a phonetics Congress like this. If you permit, I am going to start on a somewhat more general level than the one chosen by professor Fry.

May I remind those of my English colleagues whom I met at the conference I had the pleasure and the honour of attending in April of this year in Reading, of the question discussed there about the position of phonetics. Is there nowadays any reason for having special phonetic institutes and special university chairs for phonetics? Would it not be better to have just institutes of general linguistics and to look upon phonetics as a subdepartment, in some way or other, under a professor or head of department for linguistics as a whole. I think this is a serious problem in many countries. And it is not only, and not in the first place, a question of organisation and administration.

It is also a theoretical problem which concerns the place of phonetics within the wider framework of the scientific field we call general linguistics. Is phonetics a science in its own right? Can we justify a congress like this? Or would it have been better to have a section for phonetics within a congress of general linguistics as the one we had a few days ago? On the one hand phonetics belongs, in my own opinion, and must necessarily belong to general linguistics. The only motivation for phonetics as an independent science is its position as a science devoted to description and analysis of what I call the linguistic expression, with its form and with its substance, as phonemics, as instrumental phonetics, as auditory phonetics, etc. Phonetics has, and must have, its different aspects, as a study of the expression form (phonemes and prosody), of articulation, of acoustics, of perception, and so on. But all those phenomena are factors in the process of communication between people, i.e. the object of linguistics. Consequently phonetics belongs to linguistics and should be looked upon as a special field within this science. On the other hand, we all know, and we have had so many examples of this today, that phonetics has certain needs which it does not share with other fields of linguistics, needs which are partly purely economical because we need much more money than any other humanistic institute, partly related to the fact that we have to rely, for many of our problems and solutions, on collaboration with different specialists, within technique, acoustics, sound engineering, physiology, and so on. For different reasons it remains impossible, as far as I can see, to do a decent job in phonetics, if the phonetician is not allowed to concentrate on the linguistic expression and to leave other aspects of language to others. Scholars who are at the same time linguists on the content level and fully acquainted at least with the basic problems of physical analysis of sound in all its different aspects become more and more rare. The only solution of this problem I can see—and I am quite sure it will come up in my own country as well as in many others—is to do our best to conserve phonetics as an independent science with its own chairs and its own institutes though intimately connected with linguistics. This independence is motivated by its very many connections with numerous sciences outside linguistics and by its instrumental and technical needs. This is a question I think we have to face and to discuss. It is important.

In a way, this problem is intimately tied up with the problem of definitions and terminology within phonetics itself. I have been listening, yesterday and today, to different reports. And I noticed numerous differences of terminology which, I am afraid, reflect corresponding differences of concepts. It was pointed out this morning by Professor Fischer-Jørgensen that the word Phoneme was used in different ways. Some people use „phoneme“ as if it was a homonym for sound. To me and I suppose to all linguists with a structural orientation, the phoneme is not a sound. It is an abstract linguistic unit. We do not see any phonemes in a spectrogram, nor in an acoustic sound-curve. There are no boundary areas between phonemes (as they are discrete elements), just boundaries between variants or allophones (as they are manifestations of elements we call phonemes and which are linguistically not

substance units. Consequently I think we have to face the problem of our own terminology again, in order to know more exactly what we mean, and in order to avoid that one person does not talk about a sound in the same terms as another person talks about linguistic units. I know that excellent work has already been done on terminology. I am thinking for instance of Professor Hamp's vocabulary. But much more is needed. This is a problem a congress like this ought to take up again, in order to have a more clearly defined starting point for our research, for our debates and for our students.

A third question which I think must come up, is the question of a phonetic alphabet, although this is a problem which has been discussed so many times from different points of view. I am convinced that the quantitative approach to phonetics makes this problem more and more serious. If we want to use the data obtained for different languages, obtained by different research workers, in different countries, etc. we must have some kind of simplified phonetic spelling which enables us to use our data, to treat them in a modern way, statistically, mathematically, in machines, and so on. I want to propose you to take up this question again and to discuss it here before the end of this congress.

That is all I have to say for the moment.

PROF. K. PIKE (Ann Arbor):

Convictions about the future tasks grow out of beliefs about language:

1) Language is behavior—coded behavior, not an abstract code by itself. Therefore we need further study of articulatory *action*. Such study may serve as a bridge to (or from) other studies. 'Mr. John Smith', for example, is not totally an abstraction. We meet him, know him—he is not invisible. Analogously, a phoneme (or allophone) is audible.

2) Action is hierarchically organized. Therefore would study further the rhythmic organization of sounds into syllables, stress groups, etc., more intensively, in their physical and functional aspects and their acoustic correlates. Heuristic procedures for them need further attention.

3) Language can be viewed from perspectives of particle, of wave, and of field. Phonemes could be seen as normally one or the other (-consonants and vowels as particles, english stress as syntagmatic wave, english pitch levels as paradigmatic field).

Yet each phonological unit can be viewed from each view point. Phonemes, seen as arbitrarily segments, can also be treated as physical waves or as points in a field.

Perhaps the analogy of morphological fields (see articles on Fore, Potawatomi, etc.) might serve as an analogy for studying the field structure of acoustic cues.

ПРОФ. В. А. АРТЕМОВ (Москва):

Фонетика все более и более становится комплексной наукой. Она становится все более развитой и самостоятельной через смежные науки, решая их методами собственные вопросы. В таких условиях легко сбиться с пути, потерять правильную ориентировку в своем развитии. Чтоб избежать это следует возможно шире сотрудничать с смежными науками, особенно с физикой, математикой и психологией. В частности весьма существенна самая тесная связь не только с математикой, статистикой, физикой, но и с фонологией.

PROF. P. DELATTRE (Santa Barbara):

I am not so sure as Prof. Malmberg that phonetics is not a science in its own right. Some of the most brilliant phoneticians, today, are men who have come to phonetics, not by way of linguistics, but by way of other disciplines, medicine, speech pathology, physics, and above all engineering. Where would phonetics be today if Frank Cooper, Gunnar Fant, Martin Joos, Peter Denes, all engineers, had not been attracted to phonetics.

PROF. G. F. MEIER (Berlin):

Phonetiker müssen nicht Linguisten sein, aber die Sprache, die sie untersuchen, wissenschaftlich kennen. Zwischen Linguistik (Wissenschaft der langue) und Kommunikationswissenschaft der parole entwickelt sich eine Differenzierung. Beiden Zweigen kann eine besondere Phonetik untergeordnet werden. Die Wissenschaft der langue bedarf einer deskriptiven und meßtechnischen Phonetik, auch für die historische Arbeit. Die Wissenschaft der parole verlangt kollektive Forschungen der Physiologie, Pathologie, Akustik, Meßtechnik, Psychologie, Mathematik, Pädagogik, Kybernetik u. a. Es muß der gesamte Vorgang der Signalgebung von genetischer Phonetik, akustischer Phonetik (Gennematik) und funktioneller Phonetik untersucht werden. Die Vorgänge der Dekodierung verlangen psychologische Teste. — Notwendig ist dafür, Spezialisten an den Universitäten auszubilden, die alle Randgebiete in dem Maße studieren, wie sie die sprachliche Kommunikation betreffen. Wir haben ein solches Experimentalstudium begonnen.

PROF. R. GSELL (Grenoble):

Je ne voudrais ajouter que quelques mots aux exposés si riches de Mrs. Malmberg, Artemov et Pike pour dire qu'un des problèmes les plus importants de la phonétique d'aujourd'hui consiste à repenser sa situation à l'intérieur de la linguistique. Comme

un grand nombre d'études portent généralement sur la substance phonique dont la description exige des techniques de plus en plus complexes, on a trop l'impression que la phonétique est une discipline à part séparée de la linguistique. Il n'en est rien: elle en est une partie aussi ou même plus indispensable que la morpho-syntaxe ou la sémantique. L'École de Prague avait déjà rétabli l'équilibre en créant avec la 'phonologie' une phonétique fonctionnelle (à base articulatoire) qui intégrait en une unité supérieure l'étude de la substance et de la forme de l'expression linguistique. L'immensité de l'information recueillie aujourd'hui tant sur la substance phonique (physiologique et articulatoire, acoustique, perceptuelle) que sur les unités fonctionnelles et formelles du code (phonèmes en tant que constituants de morphèmes ou de monèmes et règles de construction des codes comme celles qu'étudie M. Šaumyan) nous oblige, après la linguistique de Prague, à reconsidérer ce problème. La tâche la plus urgente de la recherche phonétique me semble être actuellement l'élaboration d'une théorie générale des relations entre les différents niveaux de la substance (et de ses variations) et les différentes structurations du code qu'ils manifestent: cette tâche pourrait être assurée par une nouvelle théorie des distinctive features dont la base serait la perception. Le niveau perceptuel dont nous mesurons de plus en plus l'importance (Professeur Delattre) semble occuper actuellement (comme l'a montré Mme Čistovič) une situation centrale autour de laquelle les autres niveaux peuvent s'organiser. Le fait que la phonétique utilise des méthodes connues également dans les sciences exactes et qu'elle a souvent des applications extralinguistiques ne doit pas nous faire oublier qu'une discipline n'est pas définie par ses méthodes ou ses applications mais par l'objet de son étude: celui de la phonétique est éminemment linguistique puisqu'il concerne le plan de l'expression dans le langage, dans sa totalité.

PROF. O. S. AKHMANOVA (Moskva):

There is a common fund of knowledge which teachers of Phonetics have been spreading for years: the phoneme, distinctive features, the typology of phonological oppositions etc. form now a solid basis for much practical work in the field. The moment the student leaves the class-room, however, he is confronted with such an overwhelming variety of different 'mispronunciations' of the 'phonemes' of a language such as, say, English that he cannot fail to lose his former naïve trust in what the phonologist had tried to make him believe. One cannot, therefore, overrate the importance of the present congress where so much attention is given to the facts of human communication and such excellent and up-to-date methods have been demonstrated by means of which the 'wonder' of oral speech will, it is to be hoped, be fully accounted for at long last.

Il est délicat pour un linguiste qui n'est pas proprement phonéticien, de proposer un programme de recherches pour la phonétique. Pour illustrer la différence de réaction, à un problème particulier du linguiste et du phonéticien, je mentionnerai ici celui des traits distinctifs. Le linguiste, dans ses analyses, pour classer les phonèmes selon leurs traits pertinents se fonde sur le parallélisme des accidents subis par les phonèmes qui participent à un même trait, la 'sonorité', par exemple, sans attendre que les phonéticiens se mettent d'accord sur la nature physique réelle des traits en question. La réaction du phonéticien est naturellement de déterminer dans le détail cette nature physique. Les linguistes sont bien entendu toujours très heureux de trouver dans les résultats des recherches des phonéticiens les confirmations de leurs hypothèses.

Je me permets toutefois de reprendre certains vœux présentés précédemment à cette tribune:

Alphabet phonétique: nécessité d'assouplir le système de l'API de façon à lui faire couvrir, de façon explicite, les besoins de la transcription phonologique.

En ce qui concerne les congrès, l'exemple des deux réunions phonologiques de Vienne en 1963 et 1966 a montré qu'on peut faire du bon travail autour d'un programme suffisamment délimité. Plus le programme est spécifique et plus on a de chance d'avancer. Il serait bon, comme l'a dit M. Meier, de faire le point sur ce qu'il s'agit d'enseigner aujourd'hui aux apprentis phonéticiens. Le programme présenté par M. Meier me semble un idéal qu'il ne sera pas facile de réaliser très vite partout.

Enfin, il me paraît très utile de poursuivre les recherches visant à l'établissement d'une typologie phonologique. C'est dans ce sens qu'il faut travailler plutôt que de partir à la recherche de prétendus universaux du langage.

So wie die exakten Wissenschaften, Mathematik, Physik und Chemie, nur dadurch die Grundlagen für die moderne Technik und Industrie abgeben konnten, die die Welt verändert haben, daß die bedeutenden Forscher dieser Disziplinen nichts als Forschung und wissenschaftliche Erkenntnisse im Auge gehabt haben und im Grunde, wie es viele ausgesprochen haben, nur ihre eigene wissenschaftliche Neugierde befriedigen wollten, ohne an spätere Anwendung und praktische Auswertung zu denken, so glaube auch ich, daß den vielfachen Anwendungsmöglichkeiten der Phonetik in Sprachheilkunde und Sprechpädagogik, in Aphasieforschung, Übersetzungstechnik und Nachrichtenübermittlung am besten gedient wird, wenn man nicht diese Anwendung, sondern zunächst nur wissenschaftliche Ziele im Auge hat. In diesem Sinn ist die Phonetik die „Wissenschaft von der gesprochenen Sprache“, die mit jeder linguistischen Forschung „Hand in Hand“ gehen muß, wie Rudolph von Raumer vor 130 Jahren einmal gesagt hat. Für die gegenwärtige Phonetik ergeben sich daraus drei

Aufgaben: erstens hat die Phonetik, die der Linguistik dienen will, umfassende Archive von Tonbändern freier Unterhaltung aufzubauen, denen Archive von Tonfilmen und Röntgentonfilmen anzugliedern sind, zweitens hat die Phonetik die Aufgabe, Methoden zu entwickeln, um diese Archive so erschöpfend wie möglich auszuwerten, d. h. unter allen Gesichtspunkten, unter denen ohne diese Archive linguistisch nicht geforscht werden kann, drittens hat die Phonetik die Aufgabe, die Ergebnisse dieser Analysen der Linguistik in einer Form darzubieten, die es dem Linguisten erlaubt, diese Ergebnisse für seine linguistischen Aufgaben zu verwerten. Vor etwa 30 Jahren haben wir zu diesem Zweck die sog. Textlisten geschaffen, d. h. phonematische und auditive Texte, denen von Segment zu Segment die Daten des physikalischen Signals zugeordnet waren. Für die heute anstehenden großen Aufgaben und die dazu erforderlichen Corpora reichen diese Textlisten nicht mehr aus. Es sind vielmehr den Primärarchiven, Sekundärarchive anzugliedern, d. h. Magnetbänder, die in abrufbarer Form alle phonematischen und auditiven Daten und ihnen zugeordnet alle Daten des physikalischen Signals und nach Möglichkeit auch alle physiologischen Daten der Bewegungen des Ansatzrohres gespeichert enthalten und datenverarbeitenden Maschinen zugänglich sind. Diese Aufgaben der Phonetik sind ohne Beziehung zur Linguistik nicht lösbar, aber umgekehrt werden auch linguistische Aufgaben künftig nicht mehr lösbar sein, ohne Rücksicht auf die Realisierung der tradierten Strukturen unter psychologischen, physikalischen und physiologischen Aspekten.

Since Chao (1934) phoneticians have recognized and even enjoyed the choice offered by the non-uniqueness of phonemic solutions. But as phoneticians we have probably also all felt the need for a unification of phonemic systems for practical research and teaching purposes.

In research, we need a standardized system

1. for writing sample phonemic text in publications
 2. for establishing statistical information about language, including estimates or data-processing calculation of phonemic load in dictionary or running text, and furthermore
 3. for evaluating dialect material.
- In teaching, a unified system will be useful
4. in the arrangement of phonetic dictionaries, and
 5. in the organization of manuals of pronunciation.

I find it important that we be practical and efficient in this matter, so if we can agree, a resolution could be formed and presented during the closing session of this Congress, to recommend the establishment of an official phonemic alphabet for each language, on the pattern of official orthography.

At present phonetics is founded on, and connected with, its neighbouring branches of science, including acoustics, physiology, psychology, mathematics, electronics, sociology, pedagogy, etc. In view of this fact it is natural that the plural name 'phonetic sciences' tends to be widely used of late. In a sense, this new tendency seems to suggest something progressive; but it should be borne in mind that, despite its wide scope and ever-increasing aspects, *phonetics has its sole centre which is essentially human.*

Speech waves themselves are nothing more than a physical phenomenon which cannot be described as sound yet. It is not until one perceives them that they cause a sound; and the process of perception in this sense can be ascribed to the function of the human cerebrum. Phonetics is linguistics, but both the fields are different: 'linguistics' vs. 'phonetics' is 'static' vs. 'dynamic' and shall be called 'Science of Code' against 'Science of Speech'.

PROF. G. LINDNER (Berlin):

Meine Damen und Herren!

Es ist schon viel über die Phonetik und ihre Ziele gesagt worden. Hauptsächlich betraf dies das Wesen der Phonetik. In zweiter Linie betraf es die Forschung auf dem Gebiet der Phonetik und ihre weitere Entwicklung. Wenig ist meiner Ansicht nach bisher darüber gesagt worden, wozu eigentlich die Phonetik dient und was die Außenstehenden von der Phonetik erwarten können und dürfen. Hierbei bin ich der Ansicht, wir können die Phonetik nicht als eine „reine Wissenschaft“ auffassen, die im luftleeren Raum und isoliert von anderen Wissenschaften existiert, sondern wir müssen die Phonetik als eine Wissenschaft betrachten, die mit vielen anderen Wissenschaften Verbindungen hat und diesen durch ihre Forschungen dient. Wenn wir diesen Standpunkt akzeptieren, dann können wir nicht einfach auf das Geschaffene verweisen und die anderen auffordern, es sich anzueignen. Wir müssen vielmehr das Gedankengut, das wir im Rahmen unserer Wissenschaft aufgebaut haben, für die Anwendung in einem anderen Zusammenhang aufbereiten. Dazu gehört, daß wir uns mit den Problemen, die die anderen haben und mit ihren Forderungen an die Phonetik vertraut machen, damit wir ihnen erklären können, wie sie unsere Erkenntnisse auf ihrem Gebiet anwenden können. Das gilt beispielsweise für die Sprecherziehung im Bereich der Muttersprache. Das gilt für Fragen, die mit der zweckmäßigsten Kodierung zusammenhängen. Das gilt auch für die Defektologie sowie für den Fremdsprachenunterricht. Der Phonetiker muß sich in die angrenzenden Nachbardisziplinen einfühlen und seine Ergebnisse so weit aufbereiten können, daß sie anwendungsbereit werden. Erst damit verleihen wir der Phonetik die Ausstrahlungskraft auf andere Gebiete. Das heißt nicht, daß wir die gesamte Arbeit der Anwendung selbst leisten

sollen. Das sollen und müssen die anderen tun. Wir aber müssen den Grund legen, daß unsere Ergebnisse tatsächlich auch von den anderen genutzt werden können. Ohne diese Aufbereitung ist unsere Arbeit nur die Hälfte wert.

In diesem Zusammenhang erscheint mir eine Frage besonders wichtig. Wir als Phonetiker haben die große Aufgabe, das Mittel der menschlichen Kommunikation, die Sprache, zu erforschen, ein Mittel, das in der Alltagspraxis immer und immer wieder gebraucht wird. Ein wesentlicher Grund dafür, daß gesprochen wird, ist der, daß der Sprecher beim Hörer eine Wirkung erzielen will. Der Sprecher spricht nicht um zu sprechen, sondern er spricht, um den Hörer zu beeinflussen. Wenn wir die Anwendung phonetischer Erkenntnisse in der Praxis vorbereiten wollen, so müssen wir uns auch mit der Frage auseinandersetzen, wie das Signal, das der Sprecher produziert, auf den Hörer wirkt. Wir müssen den Sprecher befähigen, so zu sprechen, daß er beim Hörer die Wirkung erzielt, die er erzielen möchte und nicht eine Wirkung hervorruft, die er weder erwartet noch beabsichtigt. Das ist außerordentlich wichtig. Dazu müssen wir aber die Bedingungen kennen, die den Perzeptionsakt beim Hörer (beim Perzipienten) bestimmen.

Viele Untersuchungen galten den akustischen Signalen, so wie sie unter optimalen Bedingungen produziert werden. Durch diese Untersuchungen sind viele Einzelmerkmale des akustischen Signals festgestellt worden. Das akustische Signal löst beim Hörer eine bestimmte Wirkung aus. Dem steht die Tatsache gegenüber, daß in der Alltagspraxis das akustische Signal in vielen Fällen gestört ist und daß es trotzdem die beabsichtigte Wirkung beim Hörer erzielt. Das ist nur dadurch zu erklären, daß beim Hörer bestimmte Mechanismen vorhanden sind, die es gestatten, bei der Dekodierung den wesentlichen Nachrichteninhalte zu entnehmen. Über die Art und Weise, wie diese Dekodierung erfolgt, wissen wir noch verhältnismäßig wenig. Daher bin ich der Ansicht, die Phonetik sollte sich nicht nur mit ihren traditionellen Gebieten, der Produktion und der Struktur der akustischen Signale beschäftigen, sondern sie sollte auch die Problematik der Dekodierung des Signals durch den Hörer in die Forschung mit einbeziehen. Gerade weil wir auf diesem Gebiet, wie gesagt, noch wenig wissen, glaube ich, daß sich im Laufe der Zeit — vielleicht schon in nächster Zukunft — der Schwerpunkt der phonetischen Forschung auf dieses Gebiet verlagern wird. Das könnte man auch aus der Geschichte der Phonetik folgern. Zuerst stand die *Produktion* lautsprachlicher Signale im Vordergrund. Im letzten Vierteljahrhundert rückte mehr und mehr die *Struktur* des akustischen Signals in den Mittelpunkt der Forschungen. In der Zukunft werden, sollen alle Teile des Kommunikationsprozesses gleichermaßen bekannt sein, die *Vorgänge beim Perzipienten* in den Mittelpunkt rücken. Wenn diese Prognose zur Tatsache werden sollte, dann dürfen die Bedingungen des Perzeptionsprozesses nicht isoliert untersucht werden, sondern dann müßten von Anfang an gleichzeitig die Beziehungen vom Hörer einerseits zum Produktionsprozeß der Signale und andererseits zur Struktur des Signals berücksichtigt werden.

Recht schönen Dank für Ihre Aufmerksamkeit.

Monsieur le Président, Mesdames, Mesdemoiselles, Messieurs!

Dans une certaine mesure à la différence de l'avis de notre cher confrère et mon condisciple M. Martinet selon lequel plutôt les réunions phonologiques de Vienne et les congrès internationaux des linguistes seraient compétents pour traiter les problèmes phonologiques, j'insiste sur la compétence de notre congrès pour traiter les questions phonologiques non seulement en général, mais aussi en particulier en connexion avec la typologie linguistique, y compris la typologie phonologique (phonétique).

Vous savez tous que nous avons à notre disposition une typologie linguistique traditionnelle solidement fondée. M. Martinet a bien dit tout à l'heure que, dans les années dernières, on a élargi et approfondi cette typologie. Ici, à Prague — je peux à bon droit le constater — M. Skalička et d'autres linguistes ont travaillé beaucoup et avec succès dans ce domaine. Mais je propose de compléter cette typologie d'un point de vue plus synthétique et plus général selon *deux axes*: l'un horizontal *simple* — *compliqué* et l'autre vertical *régulier* — *irrégulier*. Par ex. le *français* est une langue relativement très compliquée et très irrégulière. Avec plus de détails: tandis que le système phonologique des voyelles est très compliqué, celui des consonnes est, au contraire, moins compliqué. Une variété stylistique demande le maniement différencié du phonème réduit *ə* et des phénomènes de liaison. D'autre part, il n'y a qu'un nombre très restreint des variantes combinatoires. *L'espagnol* comporte, au contraire, un système phonologique relativement simple, dans le cadre des voyelles aussi bien comme dans celui des consonnes. Mais il possède une quantité nombreuse de variantes combinatoires — même du type rare de *m* labiodental à côté de *m* labial (bilabial).

Au second degré, on doit comparer la typologie phonologique à la typologie morphologique (y compris la dérivation). En français, par ex., le système du nom est relativement moins compliqué et moins irrégulier par comparaison au système du verbe. Les alternances morphologiques y sont relativement plus nombreuses dans le système du verbe (comme en anglais, en allemand et ailleurs) qu'elles ne sont dans celui du nom. Et ainsi de suite.

Certes, tout ce que je viens de dire est suffisamment connu quant aux langues scientifiquement le mieux élaborées, quoique même ici les renseignements soient souvent très dispersés et, par conséquent, peu accessibles, surtout si l'on en a besoin à l'occasion d'un travail typologique urgent. La situation est infiniment moins favorable quant aux langues pour nous exotiques. Ni dans la 2^e édition augmentée du livre monumental et plein de mérite, intitulé *Les langues du Monde*, les caractéristiques synchroniques des langues ou des familles de langues ne sont élaborées d'un même point de vue (quelquefois, surtout dans la majorité absolue des langues américaines, elles manquent totalement).

Voilà pourquoi je propose de réaliser une enquête minimale, en adressant un *questionnaire* (Combien y-a-t-il de voyelles, sonantes, diphtongues, consonnes,

variantes stylistiques et combinatoires, types d'alternances vocaliques et consonantiques etc.; cas, nombres, personnes, modes, temps etc. Où et combien y-a-t-il d'irrégularités etc.) aux instituts et sociétés linguistiques pour ramasser et publier à bref délai les renseignements précis et utiles à confrontation avec des pourcentages concernant les caractéristiques partielles et totales relatives selon les deux axes mentionnés: *simple* — *compliqué* et *régulier* — *irrégulier*.

ГЛАСНЫЕ ЗВУКИ КИРГИЗСКОГО ЛИТЕРАТУРНОГО ЯЗЫКА

Т. К. АХМАТОВ*

Исследование звуков речи того или иного языка с помощью рентгена имеет весьма важное значение: на основе рентгенографирования устанавливаются основные артикуляционные характеристики обуславливающие различия между отдельными звуками.

В нашей работе использованы данные статистического рентгена. Как известно, статистический рентген отличается от кинорентгена тем, что он фиксирует какое-то одно положение артикулирующих органов, а кинорентген позволяет исследовать звуки в потоке речи.

Все гласные звуки киргизского языка сняты в открытых слогах после передне- и заднеязычных согласных.

В современном киргизском литературном языке имеются 14 гласных фонем (из них 8 кратких, 6 долгих). Краткие и долгие гласные отличаются между собой количественно, а не качественно.

Фонема (*a*) это наиболее открытий заднерядный сильно выдвинутый вперед, низкого подъема, неогубленный гласный.

Фонема (*a*), будучи гласной самого низкого подъема, обладает наибольшим показателем длины и высоты переднего резонатора. Масса языка оттянута назад, кончик языка находится позади нижних резцов, расстояние между корнем языка и задней спинкой зева меньше, чем при других переднерядных гласных. Однако по сравнению с заднерядными губными гласными (*u*), (*o*) больше.

При артикуляции фонемы (*a*) задняя часть спинки языка поднята к небу и высшая точка находится в ее глубокозадней части против мягкого неба, в передней части спинки языка образуется продолговатая впадина. Среди неогубленных гласных фонема (*a*) обладает самым низким характерным тоном переднего резонатора. Она в любом фонетическом положении слова произносится ясно и отчетливо. При произношении гласного (*a*) губы нейтральные.

Гласная фонема (*a*) после переднеязычных согласных имеет передне-

* Киргизгосуниверситет, кафедра киргизского языка.

рядный аллофон, т. е. язык своей передней частью сильно выдвинут вперед, а кончик языка лежит за нижними зубами на очень маленьком расстоянии. После носовых согласных (*m*), (*n*), (*ŋ*) фонема (*a*) имеет назализованный аллофон. В сочетании с билабиальными согласными лабиализуется.

При артикуляции фонемы (*i*) вся масса языка сильно продвинута вперед и приподнята. Самый высокий подъем падает на 78 перпендикуляр, кончик языка находится у десен нижних резцов, боковыми краями не соприкасается с небом, губы нейтральные.

В сочетании со среднеязычными согласными имеет наиболее закрытый аллофон. Носовые назализуют гласную (*i*).

В современном литературном киргизском языке фонема (*i*) является переднерядной слабо отодвинутой (индекс — 78), сильно прогнутой на средней части спинки и на корне языка с переходом на слабую прогнутость по всей средней и передней частям спинки языка; первой ступени подъема, неогубленная, неназализованная.

При артикуляции фонемы (*e*) передне-средняя часть спинки языка слегка приподнимается к твердому небу, кончик языка опущен вниз и только задней поверхностью касается нижних резцов. Губы при артикуляции фонемы (*e*) нейтральные, углы губ чуть растянуты. Губы слегка обнажают самые края верхних и нижних зубов.

Таким образом, фонема (*e*) является переднерядной, слабо отодвинутой, третьей приподнятой ступени подъема, с сильно прогнутой на всех участках спинки языка, включая корневую часть языка, неогубленной, неназализованной.

Фонема (*y*) является гласным смешанного ряда, выдвинутым вперед (индекс — 49), все тело языка поднято по направлению твердого неба с захватом передней части мягкого неба со слабо прогнутой на средней и задней частях спинки языка с заходом 2,3 на переднюю часть спинки языка, второй ступени подъема, неогубленный, неназализованный.

При артикуляции фонемы (*ø*) вся масса языка продвигается вперед, боковые края касаются неба, образуя значительную площадь касания по сравнению с фонемой (*ø*), кончик языка находится за нижними зубами, касаясь их нижней части и десны.

По своей конфигурации язык имеет выпуклую форму в средней части спинки языка. Губы выдвигаются вперед и округляются, т. е. образуют круглую щель.

Фонема (*ø*) — широкая гласная переднего ряда, четвертой ступени подъема, огубленная.

При артикуляции фонемы (*y*) язык сильно продвигается вперед и приподнят вверх. Конфигурация языка, начиная с задней части спинки языка, очень сходна с формой языка при артикуляции фонемы (*ø*). Кончик языка находится за нижними передними зубами и касается десен. Края

задней и средней части спинки языка касаются твердого неба. Губы сильно выпячены и округлены.

Таким образом, фонема (*y*) является закрытым звуком переднего ряда второй ступени подъема, огубленным, напряженным неназализованным гласным.

Губные гласные заднего ряда характеризуются противоположной артикуляцией языка по сравнению с губными гласными переднего ряда, при которых вся масса языка отодвигается назад, кончик языка находится в порядочном расстоянии от передних верхних и нижних резцов. Спинка языка принимает такое положение, в котором задняя часть приподнимается к мягкому небу, корень языка приближается к задней стенке зева.

При артикуляции фонемы (*u*) все тело языка значительно оттянуто вглубь ротовой полости, задняя часть спинки языка поднимается к мягкому небу, губы округляются и выдвигаются вперед. Следовательно, фонема (*u*) является огубленным, широким чистым гласным заднего ряда, второй ступени подъема.

При артикуляции фонемы (*o*) язык отодвигается назад, задняя часть спинки языка поднята к мягкому небу в большей степени, чем при (*u*). Край задней части спинки языка не соприкасается с мягким небом, кончик языка несколько заострен, собран, оттянут назад и подогнут книзу. Губы округлены больше, чем при (*u*). По всей конфигурации язык имеет выпуклую форму.

Фонема (*o*) является огубленным широким чистым гласным заднего ряда, второй ступени среднего подъема, неназализованный, напряженный звук.

Таким образом, данные статистической рентгеноъемки уточнили и поновому представили систему артикуляционной характеристики гласных современного киргизского литературного языка.

THE ANGLICIZATION OF HISPANIC PLACE NAMES IN COLORADO, USA

JOHNNYE AKIN*

The Territory of Colorado, USA, was admitted to the Union as a state in 1876. Many ethnic groups and cultural influences have played a role in the settlement and development of Colorado. The Hispanic sub-culture has been the greatest influence with the exception of the Anglo-American.

With the establishment of Mexico City as the capitol of New Spain the early Spanish explorers were able to colonize this part of the world at least a century before French explorers began to penetrate the region. After the defeat of the Spanish Armada in 1588 King Philip of Spain turned with avenging zeal to the settlement of the Pacific before the English took possession.

The early Hispanic explorers needed land marks for guidance. They had been instructed by the Spanish king to name all places, for he believed that naming was part of empire. They fought the native Indians for possession of the land and proceeded to enculturate the people and Hispanize the area.

On the Atlantic seaboard the Colonies gained independence from England in 1776 and with independence they began the settlement of the west. The English speaking settlers came late, but remained longest and in greater numbers. Although many places had been named they rejected most of them and gave new names of their choice and in their own language.

Today Colorado has a population of approximately two million, 75,000 of whom are bilingual, Spanish surnamed citizens. Because of the Hispanic subculture and the predominance of the Anglo-American linguistic influence this research was undertaken. The aim was to investigate the approximately one hundred eighty Hispanic place names that remain and to determine the extent of Anglicization used in the native culture. The conclusions are based on examples of actual usage of several hundred informants who were interviewed.

SUMMARY OF THE VOWELS

1. The vowel *a*: Initial *a* → /ə ~ æ/. *Alamosa* → /æləməʊsə/. Medial *a* → /æ ~ ə ~ ə ~ a/ depending on the amount of stress. *Baca* → /bæ ~ ka/.

* University of Denver, Denver, Colorado.

2. The vowel *e*: Initial *e* → |ɛ|. *El Paso* → |ɛl pæso|. Medial *e* → |ɛ ~ ɪ ~ i ~ eɪ|. *Del* → |dɛl|; *Dolores* → |dɔləʊri ~ ɪs|. Exceptions: *Hermosa* → |hɛmɔʊsə|; *Verde* → |vɛdɪ ~ vɛd|. Final *e* → |ɪ ~ ə|. *Grande* → |grændɪ ~ ə|. *de* → |də|. Exception: *Santa Fe* → |eɪ|, sometimes |i| thus, |sæntɪ feɪ ~ i|.

3. The vowel *i*: Initial *i* → |aɪ ~ ɪ|. *Idalia* → |aɪdeɪlɪə|. Medial *i* → |ɪ ~ i| with more frequent use of |ɪ|. *San Luis* → |sæn luɪs|; *Bovina* → |bovɪnə|. In a few words the expected |i| → |aɪ| as in *Limón* |lɪmən|.

4. The vowel *o*: Initial *o* → |oʊ ~ ɔ|. *Oro* → |óvɔro|; *Orlando* → |ɔrlændo|. Medial *o*, if it receives the primary stress, → |oʊ| as |Cebolla| |səbɔʊlə|. If it receives the weaker stress |o ~ ə| as *Bonita* |bo~nɪtə|. In a few words such as *Norte, Florida* → |ɔ ~ a|, thus |flɔ ~ arɪdə|. Final *o* unstressed usually remains |o|, less frequently |ə|. *Rico* |rɪko|.

5. The vowel *u*: Medial *u* → |ʌ ~ u ~ ə|. *Burro* → |bʌ ~ uro|; *Archuleta* → |arčə ~ ulétə|. *u* as an element of a diphthong usually remains |u|. *Pueblo* → |pjúeblo ~ píeblo ~ péblo|.

SUMMARY OF THE CONSONANTS

1. Dental |t, d| quality of Hispanic speech → English |t, d|. 2. The unaspirated quality of voiceless Hispanic |p, t, k| → English, |p^h, t^h, k^h| as *Pando* |p^hæ ~ ando; *Timpas* |t^hɪmpəs|. 3. Fricative |β, ð, ɣ| → |b, d, g| as in *Adobe* |ədɔʊbrɪ|; *Carbonera* → |k^harɔnɛrə|; *Segunde* |səgundə|. 4. Trilled |ʀ| has been abandoned for English fricative |r| as in *Rico* |rɪko|. 5. Initial *h* has become |h| as in *Hermosa* |hæmóʊsə|. 6. The palatal friction of *j* → |j, h|, less frequently |j| as in *La Junta* |la j ~ huntə| or |jantə|. 7. Palatal |λ| has been abandoned for |l, lj| less frequently |j| as in *Estella* |estelə ~ estelja| or |esteja|. 8. *s* often becomes |z| as in *Dolores* |dɔləʊri ~ ɪz|; *Kumbres* |kumbrɪs|. 9. *v* in graphemics always → |v| in the Anglicized usage as in *Vigil* |vɪjɪl|. 10. *ñ* → | ~ nj, less frequently |n| as in *Cañon*. Other consonants are treated almost exactly as the English pronunciations.

The stress pattern has shifted to the first syllable in the following words: *Conéjos* → |k^háɲɪjɔs ~ k^háɲɪhos|; *Cimarron* → |sɪmɪrən|; *Limón* → |lɪmən|; *Arriba* → |áɾɪbə|; *Florida* → |florɪdə|; *Pinón* |p^hɪnən|; *Romeo* → |róʊmto|; *Cristobal* → |k^hɪstəbəl|; *Vigil* → |vɪjɪl|.

The following words have become Anglicized in graphemics: *Arroya* → *Aroya*; *Bronco* → *Broncho*; *Cañon* → *Canon*; *Pinón* ~ *Pinon*; *Fruta* → *Fruita*; *Güero* → *Guero*.

The prosodic features of intonation, rhythm, stressing and timing were not considered in this paper. These features belong primarily to the sentence and not to the word.

A final summary seems to indicate that Hispanic influence in Colorado is one of names rather than orthoepy. The early Spanish explorers gave part of their culture

to this new land. The English speaking settlers who arrived later have taken the place names and changed them according to their linguistic patterns.

DISCUSSION

Hill:

I would merely like to add a comment since I come from an area in which similar phenomena occur. We, however, occasionally, go a step further and create pseudo Spanish names. One such is a street I pass every day—Verde Vista—evidently a translation of the next street, Greenview.

Marquardt:

The anglicization of Hispanic names is probably being paralleled in Colorado by Spanish speaking Americans in the Hispanicization of Anglic names there. To what extent do you know this to be true and what is happening to Anglic names among such speakers where it occurs?

LITERARY ARABIC X-RAY SOUND FILMS

SALMAN H. AL-ANI*

The material on these films (numbers 621-1 and 621-2) has been recorded by a native speaker of Iraq at the Department of Otolaryngology and Maxillofacial Surgery, University of Iowa, Iowa City, Iowa,¹ where the films are on file together with an accompanying script.

Film number 621-1 contains the following:

1. Vowels in isolation

<i>Short</i>	<i>Long</i>
i	ii
u	uu
a	aa

2. Plain vs. emphatic consonants

t	vs.	t̤
d	vs.	d̤ ²
s	vs.	s̤
ð	vs.	ð̤

These are recorded in consonants-vowel sequences and then as lexical items in minimal pairs in initial, medial and final positions and geminated.

Film number 621-2 includes the back consonants

ʔ	vs.	ε
ħ	vs.	ħ̤
ħ̤	vs.	ε
k	vs.	q
x	vs.	γ

* Institute of Islamic Studies, McGill University, Montreal, Quebec, Canada.

¹ The author would like especially to acknowledge Professors Kenneth Moll and James Lubker who assisted in the preparation of the films.

² [d̤] does not occur as a natural phoneme in my speech, either on the literary or dialectal levels. It is included here merely for the completion of the emphatic consonants on the literary level.

These are recorded in the same manner as the consonants in film number 621-1.

The equipment used in the filming of the material was designed by the North American Phillips Company. It consists of the following basic components: a Rotalix 0—75/125 X-ray tube with a 0.6 mm² focal spot, a 300 ma generator, a smoothing capacitor, and a nine-inch image intensifier tube with an intensification factor of approximately 3000. The light emitted from the intensifier tube may be directed through a lens system to either a fluoroscopic viewer or to the lens of a motion picture camera.

The films are very clear and cover the laryngeal, pharyngeal and vocal cavities. For synchronization purposes, the sound that accompanies the films is about 26 frames ahead of the picture. It takes about five minutes to project both films continuously. At the beginning of film 621-1, when projected, black dots appear on the front part of the mouth—above the tongue. Each dot measures about 1/4 inch in diameter. These dots may be used to measure the actual size of the speaker's mouth. The tongue was painted with barium, as far back as possible, in order to show the movements clearly.

The films may be ordered from the University of Iowa, either separately or together, by writing to the Department of Otolaryngology and Maxillofacial Surgery.

DIE VIBRATORISCHE KOMMUNIKATION VON *UCA TANGERI* UND *UCA INAEQUALIS*

Ein Vergleich

R. ALTEVOGT*

Neben der namengebenden Gestik des visuellen Kommunikationsbereiches (1, 2, 5) besitzen wohl alle Winkerkrabben der Gattung *Uca* ein vibratorisches Nachrichtsmittel, das rein phänomenologisch an einer Art (*pugilator*) schon früh bekannt wurde (8,7), dessen Umfang und Funktion aber erst in jüngerer Zeit mit modernen Registriergeräten erfaßt werden (3—6, 19, 10, 12, 13). Bei den hier geschilderten Befunden erfolgte das mit Beschleunigungsaufnehmer 4308 und Schallpegelmesser 2203 von Brüel u. Kjaer sowie den Batterietonbandgeräten TK 4 und 6 von Grundig.

Bei *U. tangeri* stellte sich dabei heraus, daß es hier zwei Arten vibratorischer Äußerungen gibt: den kurzen, allenfalls dreisilbigen Wirbel und den langen Wirbel mit max. 12 Impuls-, „Silben“. Beide werden durch Schläge der Winkschere gegen den Boden bzw. gegen die oder eines der Gehbeine produziert. Dabei ist die Homologie von Vibrations- und Winkkommunikation evident: Der kurze Wirbel entspricht mit Abständen von durchschnittlich 1,1 s dem Winken der „Ruheaktivität“ und signalisiert lediglich die Anwesenheit eines kopulabereiten Männchens. Er geht in den langen Wirbel über, sobald oberirdische Schrittvibrationen die Annäherung potentieller Sexualpartner oder Höhlenkonkurrenten ankündigen.

Dabei bewirken Schrittnäherung (als Schallpegelsteigerung) und intermittierendes Verharren des schreitenden Krebses Steigerung der Impulsanzahl pro Wirbel und Verkürzung der Wirbel- und Impulsintervalle. Die Wirbelabstände verkürzen sich von 1,5 s auf 0,8 s, und die Impulsintervalle innerhalb des kurzen Wirbels sinken von 0,10 auf 0,08 s, während sie im langen Wirbel recht konstant auf 0,16 s gehalten werden. Daraus folgt, daß die Information „Hier ist ein kopulabereites ♂“ (beim kurzen Wirbel) bzw. „ein solches mit Höhlenbesitz“ (beim langen) wohl nicht im Impulsabstand, sondern eher in der Impulsanzahl pro Wirbel enthalten ist. Auch der Wirbelabstand dürfte Informationswert besitzen, wenngleich es schwierig ist, hier Weiteres auszusagen, da er durch Temperatur und Erregungsgrad beeinflußt wird (10).

Daß die oben gegebene Deutung des langen Wirbels richtig ist, läßt sich durch einen Klopfkode zeigen, den wir 1963 bei *U. tangeri* in Marokko ausarbeiteten.

* Abteilung Physiologie und Ökologie des Zoologischen Institutes der Universität Münster.

Klopft man bei balzlustigen *tangeri*-♂♂ mit richtiger Impulsstärke (Schalldruck) im (etwa) richtigen, d. h. arttypischen Impulsabstand mit dem Finger auf den Boden, so kann man die ♂♂ in den Höhlen zum vibratorischen Antworten bringen (4, 6). Schließlich steigen sie im Laufe solcher Dialoge aus den Höhlen, um sich dem vermeintlichen Kontrahenten zum Kampf um die Höhle zu stellen. Bei Weibchen gelingt dieses Emporlocken der sonst so scheuen Tiere ebenso, sie „erwarten“ ja oben den Kopulapartner und lassen durch den klopfenden Menschen das Kopulavorspiel mit Beinzerren und Carapax-feeding über sich ergehen (4, 6). Versuche mit Variationen solcher Klopfkodes ergaben für *U. tangeri*, daß vor allem Impulsanzahl, Schallpegel und Wirbelabstand, nicht aber die Sendefrequenz (nach Sonagrammen von 200–3000 Hz) informationsträchtig sind.

Vergleichsuntersuchungen an sympatrischen *Uca*-Arten Mittel- und Südamerikas im Herbst 1966,¹ von denen hier nur *U. inaequalis* erwähnt wird, erbrachten eine Bestätigung dieser Deutung auch für diese Art. Auch hier konnten wir einen arttypischen Klopfkode ausarbeiten, der aber mit deutlich kleineren Impulsabständen und größeren Impulsanzahlen pro Wirbel einhergehen muß, soll er erfolgreich sein (Abb. 2). Das zeigt noch einmal, daß die Impulsanzahl pro Wirbel offenbar arttypisch ist — schon in den *tangeri*-Populationen Südspaniens und Südwestmarokkos offenbaren sich Dialektunterschiede, indem man mit demselben Klopfkode in den genannten Populationen unterschiedlichen Erfolg hat, seine Verständlichkeit im N und S also wohl verschieden ist.

Mit der von *tangeri* verschiedenen Impulsanzahl pro Wirbel ist bei *U. inaequalis* auch der Impulsabstand verschieden: Er beträgt durchschnittlich 0,064 s (mit seltener Streuung bis 0,071 s). Daß er eine noch größere Streuung wahrscheinlich ertragen würde, ohne an artspezifischer Verständlichkeit einzubüßen, zeigt die Tatsache, daß das menschliche Signal sogar Impulsabstände von 0,086 s aufweisen darf und dennoch vom Krebs beantwortet wird (Abb. 3). Die Latenz der *inaequalis*-Antwort auf solche menschlichen Vibrationsanrufe beträgt durchschnittlich 2 s (min. 1,5; max. 4 s).

Daß wie bei *U. tangeri* so auch bei *U. inaequalis* jeder Winkakt von vibratorischen Impulsen begleitet ist (4), sei abschließend erwähnt und gilt u. a. auch für die ebenfalls südamerikanischen *U. mertensi*, *U. batuenta* und *U. terpsichores*.

LITERATUR

1. Altevogt, R.: Untersuchungen zur Biologie, Ökologie und Physiologie indischer Winkerkrabben. *Z. Morph. Ökol. Tiere* 46, 1–110, 1957.
2. —: Ökologische und ethologische Studien an Europas einziger Winkerkrabbe *Uca tangeri* Eydoux. *Ibid.* 48, 123–146, 1959.

3. —: Akustische Epiphänomene im Sozialverhalten von *Uca tangeri*. *Verhandl. d. Deutsch. Zool. Ges.* Wien 1962, 309–315, 1963.
4. —: Ein antiphoner Klopfkode und eine neue Winkfunktion bei *Uca tangeri*. *Naturw.* 51, 644–645, 1964.
5. —: *Uca tangeri* (Ocypodidae): Klopfen und Winken. Film E 693 der Encyclopaedia cinematographica, Göttingen 1964.
6. —: Vibration als semantisches Mittel bei Crustaceen. *Wissenschaftl. Z. d. Karl-Marx-Universität Leipzig* 15, 471–476, 1966.
7. Burkenroad, M. D.: Production of sound by the fiddler crab, *Uca pugilator* BOSCH, with remarks on its nocturnal and mating behavior. *Ecology* 28, 458–461, 1947.
8. Dembowski, J.: On the "speech" of the fiddler crab, *Uca pugilator*. *Trav. Inst. Nenci* 3, 1–7, 1925.
9. Hagen, H. O. von: Nächtliche Aktivität von *Uca tangeri* in Südspanien. *Naturw.* 48, 140, 1961.
10. —: Freilandstudien zur Sexual- und Fortpflanzungsbiologie von *Uca tangeri* in Andalusien. *Z. Morph. Ökol. Tiere* 51, 611–725, 1962.
11. —: Studien an peruanischen Winkerkrabben. *Zool. Jahrb., Syst.* 95, 395–463, 1968.
12. Salmon, M.: Waving display and sound production in the courtship behavior of *Uca pugilator*, with comparisons to *U. minax* and *U. pugnax*. *Zoologica* (N. Y.) 123–149, 1965.
13. — and J. F. Stout: Sexual discrimination and sound production in *Uca pugilator* Bosch. *Zoologica* (N. Y.) 47, 15–20, 1962.

¹ Dank freundlicher Hilfe der Deutschen Forschungsgemeinschaft, der Deutschen Ibero-Amerika-Stiftung, der Gesellschaft zur Förderung der Westfälischen Wilhelms-Universität und des Instituts für den wissenschaftlichen Film Göttingen.

Abbreviatur: Die vibratorische Kommunikation von *Uta tangeri* und *Uta inaequalis*

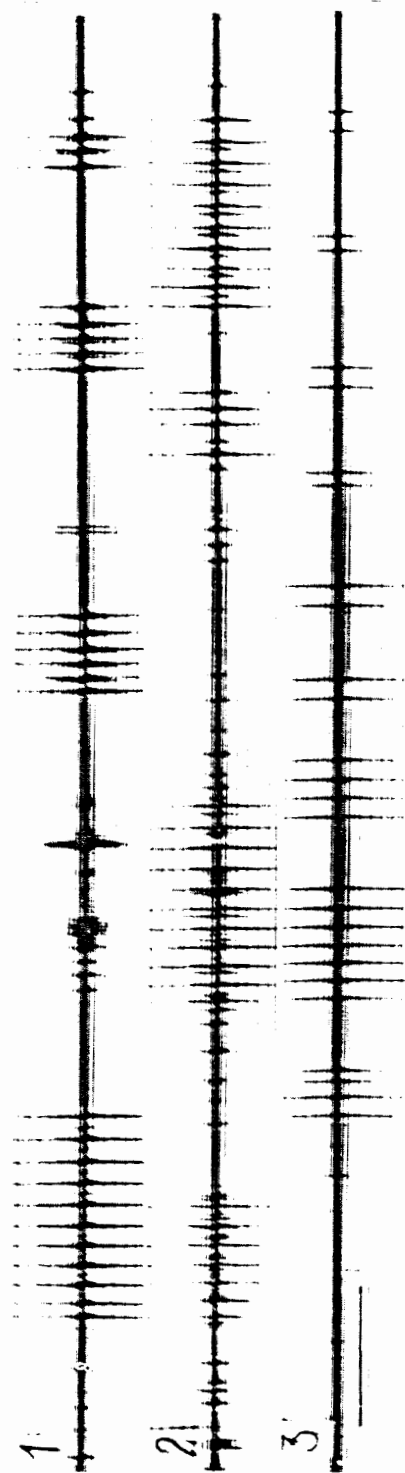


Abb. 1. *Uta tangeri*: Reaktionen der Männchen auf Schritt- und Schrittschwingungen der Weibchen. Kleine, unregelmäßige Amplituden; Weibchen; große, regelmäßige Amplituden; lange Würfel der Männchen. Zeitmarke links unten (1 sec (aus 3)).

Altevogt: Die vibratorische Kommunikation von *Uca tangeri* und *Uca inaequalis*

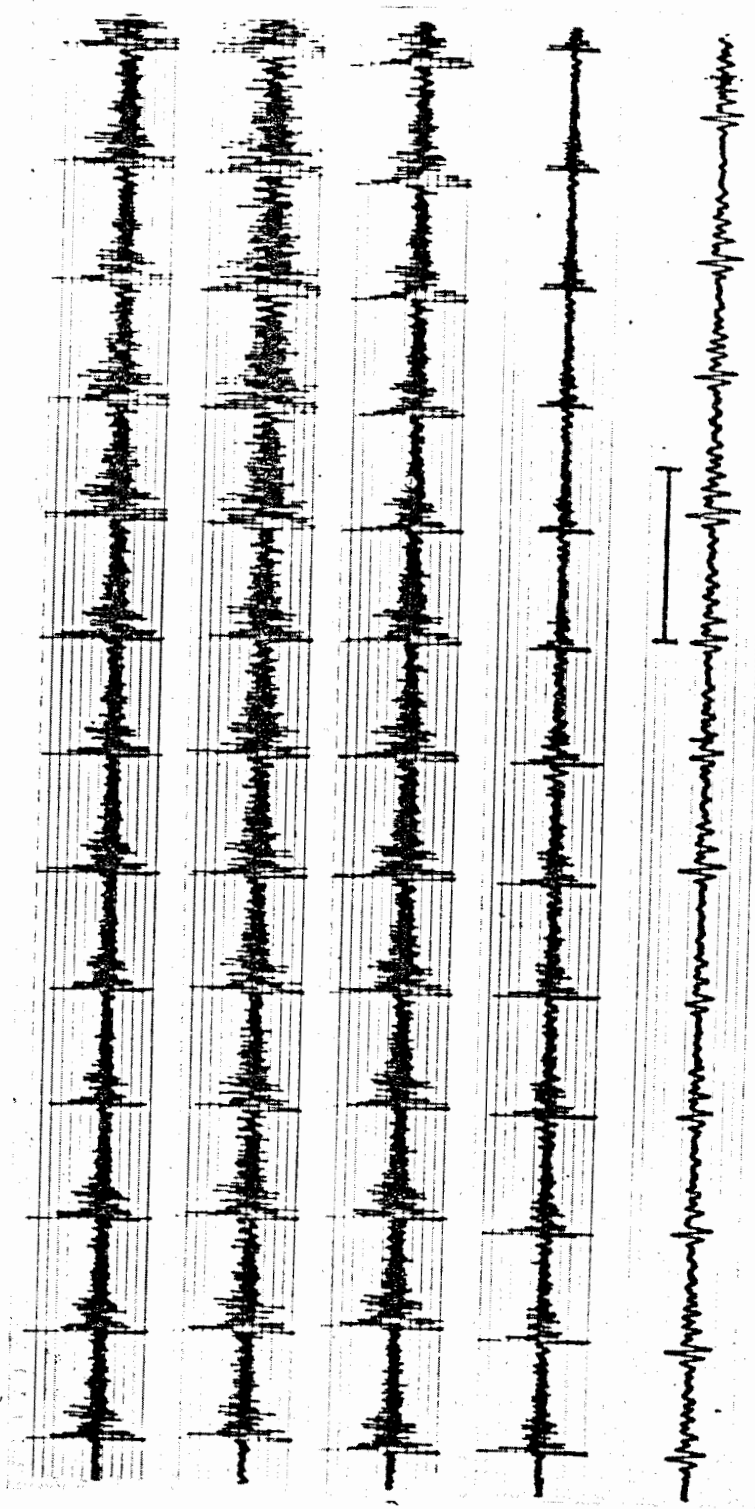


Abb. 2. Winkerkrabbe *Uca inaequalis*: Vibratorische Kommunikation der Männchen mit Artgenossen (Zeilen 1—4) und antiphones Klopfsignal des Menschen (Zeile 5), das das artspezifische inaequalis-Klopfen provoziert. Tonbandoszillogramm. Zeitmarke 0,1 sec.

Altevogt: Die vibratorische Kommunikation von *Uca tangeri* und *Uca inaequalis*

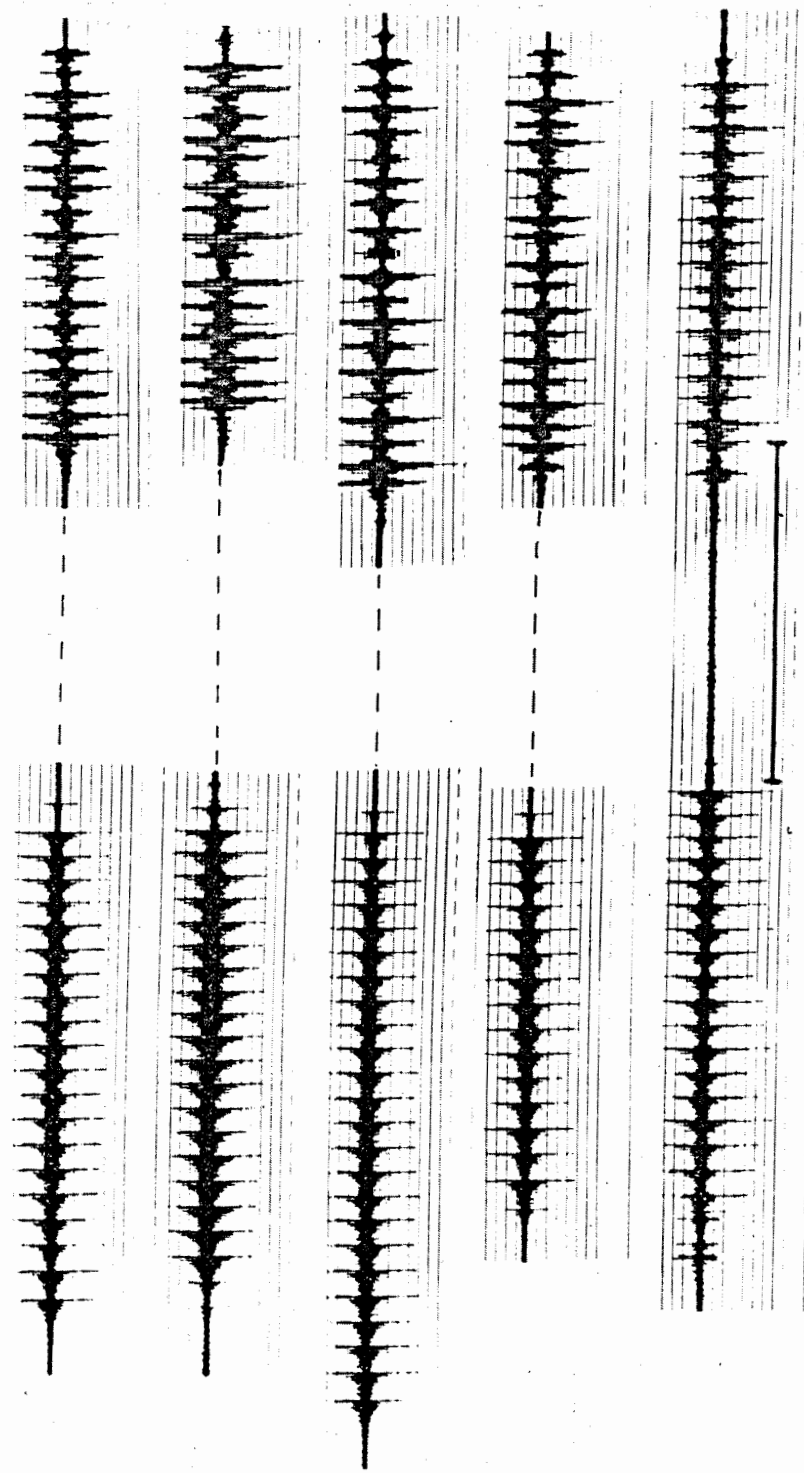


Abb. 3. Antiphones Klopfen Mensch (jeweils linke Impulsgruppe jeder Zeile) und Winkerkrabbe *Uca inaequalis* (die rechten Impulsgruppen jeder Zeile) in Puerto Pizarro (Peru) Sept. 1966. Zeitmarke 1 sec.

A COMPARATIVE STUDY OF THE VOWELS OF DIFFERENT ACCENTS OF ENGLISH

W. R. BRIAN ANNAN

This paper is a preliminary statement of my aims and method. The intention is to make a spectrographic analysis of recordings made by native English speakers, the speakers being from any area where English is spoken as an LI (e.g. R. P., General American, Australian, Scottish and Yorkshire). They must be educated to the level of University Entrance and be "mutually intelligible" and not be capable of classification as dialect speakers where by dialect I understand a form of language differing from the Standard in Lexis, Syntax and Pronunciation.

The material for recording is a list of English monosyllables read as citation forms and on a falling intonation. The spectrograms are made on a Kay Sonagraph and are on a scale of 0—4000 c.p.s. with a Broad Band resolution. Sections are taken for each vowel at an arbitrarily selected mid-point in terms of duration and values are read off at the central point of frequency of formant pattern. The resultant

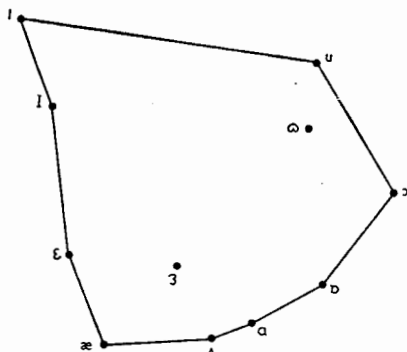


Fig. 1. Vowels of R. P. English.

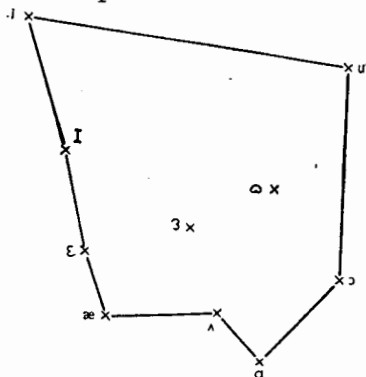


Fig. 2. Vowels of General American.

readings are plotted on a formant chart (F_1 , F_2). No account is taken of F_3 and no attempt to find average values for speakers of the "same" accent are made. The formant diagrams are copied onto tracing paper with no record of their frequencies and are compared. The method of comparison is to assume the vowel [i] as common and then to rotate the diagram in such a way as to determine which vowels are located in the same area (see Fig. 3).

T. Hill makes the statement that "we may suppose that, for complete mutual comprehension, any two speakers of a koine must have sound systems such that

'corresponding' vowel phonemes, for instance, occupy corresponding places when plotted within their systems on a cardinal vowel chart or formant chart."¹ This is the starting point of my work and has, as you will see on Fig. 3, produced some interesting points. Between two speakers, one of General American and the other an R. P. speaker, we have four points of congruity or near congruity: [ɛ] [æ] [ʌ] and [u].

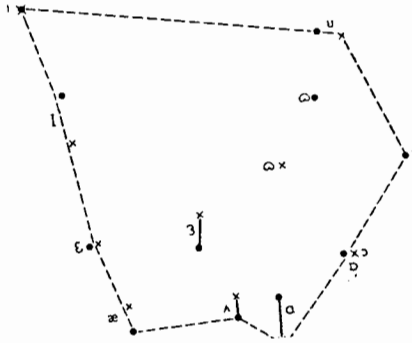


Fig. 3. Composite Picture.

These are, for me, points of mutual intelligibility within my framework. Now, if you compare R. P. [ɒ] and G. A. [ɔ] you will see congruity as well. Their functions are phonologically distinct within their respective accents and I would suspect that this is an area where 'complete mutual comprehension' would not exist, remembering that I am dealing purely with citation forms and that there is a need for other contextual clues to perception.

One of the major difficulties with which I am dealing is the problem of a vowel (monophthong) in one accent, being in systemic terms a diphthong in another (e.g. Scots [e] in 'day' as opposed to R. P. [er]). I have not included these in my example but would propose that comparison may be made by using the prolongable element of the diphthongs as my norm. Further to this point is the question of [r]-pronouncing accents; i.e. accents of English like Scots which pronounce [r] in post-vocalic position. It has been suggested that they might be treated as special types of diphthong.

In this paper I have attempted to produce little more than a tentative statement and, also, to ask some questions which I consider relevant to the problem of perception and analysis.

¹ T. Hill, 'Institutional Linguistics', *Orbis*, vol. VII, No. 2, 1958, p. 454.

ИНТОНАЦИОННАЯ СТРУКТУРА ВОСКЛИЦАТЕЛЬНЫХ ПРЕДЛОЖЕНИЙ В ЕЕ ВЗАИМОДЕЙСТВИИ С ИХ ЛЕКСИКО-ГРАММАТИЧЕСКОЙ СТРУКТУРОЙ

(на материале английского и русского языков)

А. М. АНТИШОВА*

Восклицательные предложения рассматриваются как самостоятельный коммуникативный тип предложений, в которых наряду с сообщением определенного факта выражается особое эмоциональное отношение говорящего к действительности.

Аудирование восклицательного предложения происходит на основе его лексико-грамматической и интонационной структуры. При этом под интонацией понимается совокупность мелодического, силового, темпорального и тембрального компонентов.

Восклицательные предложения в современном английском и русском языках характеризуются особой синтаксической структурой (What a nice girl she is! Какая она хорошая девушка!), а также структурой повествовательного и вопросительного предложений. Однако, какой бы синтаксической структурой не обладали восклицательные предложения как в английском, так и в русском языках, все они характеризуются особым составом интонационных признаков (7 признаков в обоих языках), которые не присущи другим коммуникативным типам предложений и которые в своей совокупности образуют своеобразную интонемику восклицательных предложений.

Следующие различительные признаки указывают на черты сходства интонационной структуры восклицательных предложений в английском и русском языках:

1. смещение частотного уровня вверх или вниз от среднего уровня, типичного для невосклицательных предложений;
2. более крутое нисходящее или восходящее завершение, воспринимаемое на слух как более резкое, в направлении основного тона, обусловленное смысловым содержанием предложения;
3. увеличение интенсивности по сравнению с интенсивностью невосклицательных предложений, что, обычно, воспринимается как увеличение громкости в восклицаниях;

* ЛЭФИНПР 1 МГПИИЯ им. М. Горького, Москва.

4. увеличение времени звучания, что воспринимается как замедление темпа;
5. яркая тембральная (эмоциональная) окраска и
6. децентрализованное ударение (все ударные слова выделены в равной степени).

Однако, было бы ошибочным поставить знак равенства между перечисленными признаками восклицательных предложений в английском языке и русском языке. Можно говорить лишь о сходных соотношениях интонационных составляющих. Поскольку в этом соотношении обнаруживается много сходных черт, то это дает основание предполагать, что в восклицательных предложениях находит выражение определенная общечеловеческая коммуникативная и психологическая категория.

Основное различие между английскими и русскими восклицательными предложениями заключено в различии их мелодического рисунка. Английские восклицательные предложения характеризуются нисходящими или восходящими мелодическим движением в шкале. Для русских восклицательных предложений типична монотонно-вокалическая шкала. (Ударные гласные располагаются на одном частотном уровне, что воспринимается как монотон.)

В связи с различным характером мелодического движения различно и взаимодействие завершения и шкалы в сравниваемых языках. В русском языке начало нисходящего завершения в восклицательных предложениях находится на одном уровне с ударными гласными в шкале, в английском языке при нисходящей шкале начало завершения находится ниже ударных гласных в шкале, а при восходящей шкале выше.

Интонационная структура восклицательных предложений находится в тесной взаимосвязи с их лексико-грамматической структурой. Если грамматическая структура или лексический состав, или то и другое вместе несут на себе часть эмоциональной нагрузки, то роль интонации в оформлении восклицательного предложения несколько снижается, и нередки случаи, когда из состава интоны выпадают один или два признака одновременно.

Такое взаимодействие грамматической структуры, лексического состава и интонационной структуры в восклицательных предложениях еще подтверждается закон „Экономии сил“, выдвинутый А. М. Пешковским (Пешковский А. М. Русский синтаксис в научном освещении, М., 1938, стр. 75). Согласно этому закону, одни языковые средства компенсируются другими, служащими той же цели.

Восклицательные предложения в современном английском и русском языках характеризуются определенной интонационной и лексико-грамматической структурой и обладают определенным коммуникативным содер-

жением присущим всей группе восклицательных предложений. Следовательно, восклицательные предложения в современном английском и русском языках являются самостоятельным коммуникативным типом.

DISCUSSION

Pike:

In support of the speaker's comment a moment ago: any set of contrasts which I can make in ordinary speech I can, I believe, also make in a very high-pitched voice, or low-pitched. The contrasts must not be tied to any absolute high or low.

Stojanović:

If a comparison between two languages is aimed at, then examples should be provided so that distinctions in pitch ranges of the two languages can be clearly seen. The pitch ranges could be determined in musical terms, or by any distinctive system.

Багмут:

1. Одной из отличительных характеристик восклицательных предложений является восходящее движение кривой основного тона на конце фразы, как это и отмечается в докладе. Меня интересует соотношение между движением основного тона и интенсивностью произнесения фразы. Существует ли прямая зависимость между этими характеристиками, т. е. возрастает ли в восклицательных предложениях интенсивность при увеличении частоты основного тона? Если возрастает, то в каком соотношении?

2. Восходящее движение основного тона, как известно, может характеризоваться как резко восходящее и восходящее медленно. Указанное Вами восходящее движение основного тона имеет большой или незначительный интервал?

Артемов:

Интонационные исследования привлекают все больше число фонетистов во всех странах, особенно в нашей стране. Назрела необходимость или включить этот вопрос в повестку следующего конгресса или созвать симпозиум между 6-ым и 7-ым конгрессами.

THE INTELLIGIBILITY OF THE SENSITIZED SPEECH TESTS IN THE DISORDERS OF THE CENTRAL AUDITORY MECHANISMS

A. ANTONELLI*

The hearing symptoms associated with disorders of the central auditory mechanisms are very often absent and sometimes of an extremely elusive nature. In a few cases only they are as handicapping as normal deafness due to cochlear or auditory nerve impairments.

This fact may help us in explaining why the first really significant breakthrough on this problem was achieved in the early 1950's. As a matter of fact the routine auditory tests—such as the pure tone threshold, the supraliminal tests with pure tones, the speech articulation curve—, failed for a long time to give reliable results.

Thus, in the course of about ten years, experience has allowed us to work out a series of more elaborate tests, some of which make use of pure tones presented under special conditions, while others use specially devised or processed speech material. The experience of our Department in this particular field is based mainly on the use of the so-called sensitized speech tests, consisting primarily of methods called “the distorted voice test”, and “the time-compressed voice test”.

We have performed our sensitized speech tests on a number of patients with spontaneous lesions, tumours, vascular disorders, inflammatory processes, multiple sclerosis etc. — of the cerebral hemispheres or the brain-stem or who had undergone surgical removal of the temporal lobe for focal epilepsy.

In spontaneous pathology, on account of the impossibility of establishing the microscopical limits of the disorders, the presumed site of the lesion was determined on the basis of an extensive clinical neurological examination.

Let us now describe separately the audiometric findings in brain and in brain-stem pathology.

For the former purpose we shall confine our remarks to those patients who underwent surgical removal of brain tissue. As compared to data established in spontaneous pathology the study of auditory function following temporal lobectomy undoubtedly offers the advantage of correlating the qualitative and quantitative aspects of the hearing performance with the anatomical limits of the lesion, based on the neurosurgical operative report.

* From the E. N. T. Department of the University of Ferrara — Italy — (Head: Prof. E. Bocca).

And this (Fig. 1) is a typical audiometric picture after right temporal lobectomy: threshold acuity for pure tones measured on both ears is well within normal limits. No patient shows any kind of significant hearing-loss, in the sense of diminished threshold acuity.

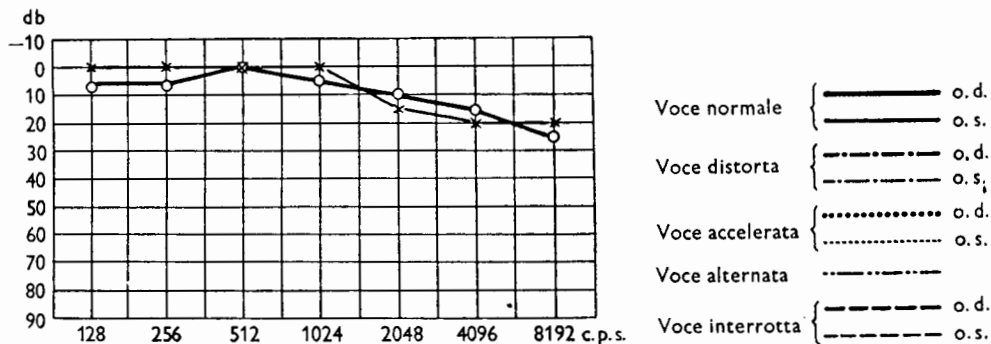


Fig. 1. Lobectomia temporale Dx.

The routine tonal tests—such as SISIGram, auditory adaptation and so on—give normal values. Intelligibility for normal speech material is in accordance with pure tone threshold acuity.

When in those patients the sensitized speech material is monaurally delivered, at progressively increasing intensities, it becomes very clear that the articulation

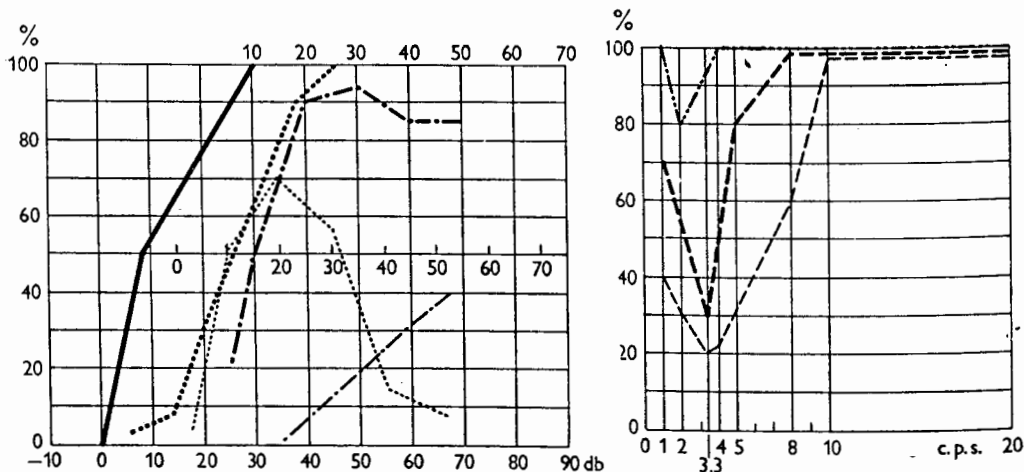


Fig. 2. Lobectomia temporale Dx.

curve is distinctly worse in the contralateral ear, where the intelligibility falls to about 50%. The ipsilateral ear gives, on the contrary, normal results (Fig. 2).

We should like to stress here the most interesting finding, i.e. that auditory function has been found to be impaired to the same extent as after the surgical

Antonelli: The Intelligibility of the Sensitized Speech Tests in the Disorders of the Central Auditory Mechanisms

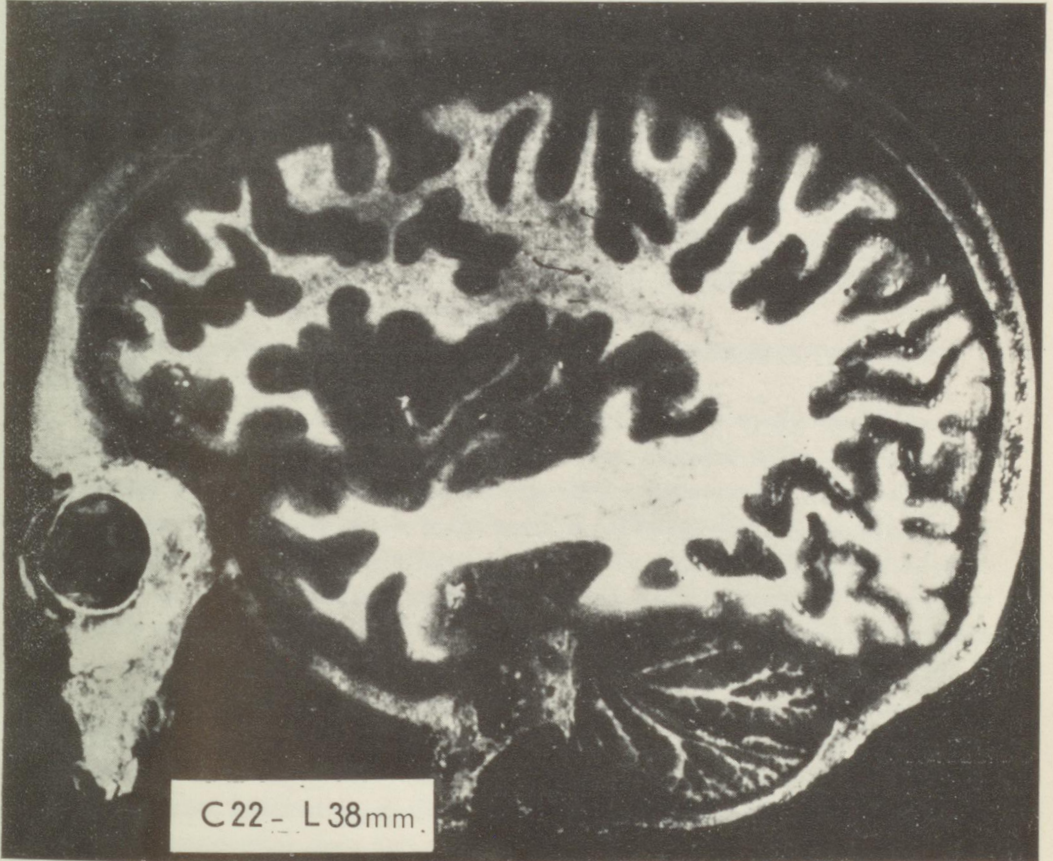


Fig. 3.

removal of the so called Heschl's gyrus, considered to be the site of the primary cortical auditory area (Fig. 3).

On the other hand the intelligibility defect regresses sooner after surgery, whenever Heschl's gyrus is spared, unless its function was impaired even before surgery.

In the disorders of the central auditory mechanisms at a brain-stem level the pure tone threshold acuity is often impaired. There is no characteristic pattern, but different aspects of high-frequency type neural loss. Poor articulation curves for sensitized speech material, with reduced intelligibility are commonly found, but—at variance with cortical dysacusis—the auditory disorder is often bilateral. This fact might be attributed to the much smaller distance separating the two auditory pathways at the brain-stem level, so that equally extended lesions yield a much higher incidence of severe and bilateral impairments when they are located in the brain-stem, than when they are located at a cortical level.

The more impaired ear has been found to be ipsi- or contralateral to the side of the brain-stem lesion in an equal number of cases.

In some patients, with little or questionable alteration of the pure tone threshold acuity a poor discrimination for normal speech is found, affecting one ear only. Such an observation is quite interesting, since it resembles a kind of unilateral "verbal deafness". We have called it "tonovocal dissociation".

Let us finally summarize the main differences between audiometric pictures in cortical and brain-stem (b.s.) pathology.

1. First of all the hearing impairment is slighter and strictly monolateral in cortical lesions, whereas it is more marked and often bilateral in b.s. lesions.

2. Another difference concerns the correspondence between the ear affected and the side of the neural lesion: while in the cortical pathology the hearing impairment is always crossed—namely it is always found on the opposite side of the affected hemisphere—in the b.s. lesions the hearing impairment is either ipsi—or contralateral with the same incidence. When tone speech dissociation occurs it is often contralateral.

3. A third difference is the occurrence of pure tone threshold impairments, which is more frequent in b.s. lesions. Although all the elements of a complete audiometric investigation lead to a general diagnosis of lesion of the central auditory paths in the b.s., the exact localisation of the disease can be established only with the aid of the other neurological signs and symptoms.

DAS PHÄNOMEN DER LAUTBEDEUTSAMKEIT AUS BIOPHONETISCHER SICHT

HERBERT ARNDORFER

Perzeption und Produktion des akustischen Sprachsignals stellen eine Wechselbeziehung zwischen Sprecher und Hörer her, deren emotionale Färbung sich nach Trojan auf physiologische Generatoren zurückführen läßt. Die Lautgestaltung emotional durchdrungener Kommunikationsformen ist ein Teilelement des überaus komplexen Gebietes des stimmlichen Ausdrucks und damit ein Anliegen der ganzheitlich orientierten Biophonetik.

„Zerhauene Waffen“ hat W. Schneider sehr treffend die unglaubliche Verwirrung der Terminologie genannt, die sich aus den nur geisteswissenschaftlich orientierten Deutungsversuchen ergab, wie etwa die seit Platons „Kratylos“ wiederholt unternommene Verknüpfung des Problems mit Theorien der Sprachentstehung, der Glaube an eine Inhaerenz der Bedeutung oder das Ausgehen von Einzellauten. Der Einzellaut im Wort kann nicht Träger einer Bedeutung sein, da er nur Baustein eines Zeichens ist (Trojan). Die Lautbedeutsamkeit — oder noch unverfänglicher: Ausdruckswirkung — kann nur von einer Häufung gleichartiger Laute ausgehen, und dies nur dann, wenn Sinn und Gefühlsgehalt übereinstimmen. Es handelt sich hier um ein echtes phonetisches Problem, bei dem der auditiv-phänomenologische Aspekt gleichberechtigt neben den artikulatorischen tritt.

Bereits in einer frühen Arbeit wurden von F. Trojan die Grundzüge dieser biophonetischen Betrachtungsweise dargelegt, als der Autor an Hand klassischer deutscher Lyrik die Abhängigkeit des Gehalts von der Verwendung bestimmter Laute, aber auch deren polare Erscheinungsformen (wie etwa das „makabre a“) aufzeigte und auf ihre vegetativen Grundlagen hinwies.

Höffe verdanken wir die experimentelle Feststellung, daß sich die Lautgestalten im Ausdruck der Emotionen verändern können.

In der bisher umfangreichsten Untersuchung von Gedichten und Balladen deutscher Dichter konnte E. Arndorfer den Nachweis eines Zusammenhanges zwischen Häufungen gleichartiger Laute und emotionalem Gehalt erbringen. Der Verfasser selbst verwendet seit geraumer Zeit Material dieser Art mit sehr befriedigendem Erfolg in der Sprecherziehung. In sprachpsychologischen Testreihen offenbarte sich eine hohe internationale Gleichartigkeit der durch Lauthäufungen bewirkten „Gefühlsresonanzen“.

Die in den erwähnten Untersuchungen bisher gewonnenen Erkenntnisse lassen sich wie folgt darstellen:

1. Prinzipiell ist zwischen „subjektiver“ und „objektiver“ Lautbedeutsamkeit zu unterscheiden. In ersterer manifestieren sich psychische Zustände, insbesondere in der Lyrik, und sie wäre etwa Kaysers Begriff der „Lautsymbolik“ gleichzusetzen; die objektive Lautbedeutsamkeit gibt äußere Vorgänge, besonders in Interjektionen, wieder und findet sich vor allem in der Ballade. In der Alltagsrede ist sie durch Aktivverben und Interjektionen vertreten.

2. Die Polarität lautlicher Ausdruckswirkung, von Werner als „dynamischer Charakter“ bezeichnet, wird von Trojan auf die vegetative Steuerung zurückgeführt und durch diese und die jeweils verwendete faukale Distanz erklärt. Während etwa im „Erkönig“ das a mit Rachenweite echte Beruhigung ausströmt, manifestiert sich in den faukal engen a-Lauten der Leichenszene in Bürgers „Lenore“ die grauenerfüllte Ruhe des Todes. Die gleiche Erscheinung läßt sich auch anderwärts nachweisen, etwa bei ei (Freude und Lockung gegenüber Leid und Jammer) oder au (Gemütlichkeit gegenüber Melancholie und Schwermut).

3. Die „Kardinalvokale“ a-i-u erweisen sich auch in ihrer Ausdruckswirkung als Superstrukturen der Register nach Trojan. Zu verweisen wäre in diesem Zusammenhang auf das „sakrale u“ der indogermanischen Frühzeit und auf die bemerkenswert subtilere Nuancierung der mittleren Vokale o und e in der Moderne.

4. Träger gefühlvollen Ausdrucks unter den Vokalen sind die Längen, während die Kürzen die der Erregung sind und damit die Konsonanten stützen.

5. Vokale und Konsonanten (die Beibehaltung dieses Einteilungsschemas sei hier gestattet) stehen einander gegenüber als Manifestationen von Vagotonie, Schonstimme und Legato einerseits und von Sympathikotonie, Kraftstimme und Stakkato andererseits.

6. Unter den Konsonanten ist ebenfalls eine Differenzierung bemerkbar: In den Plosiven p, t, k und in der Liquida r manifestieren sich Emotionen der ergotropen Funktionsrichtung, wie Kampf und Arbeit, während die Frikative und das l trophotrope Zustände, aber auch natürliche Bewegungen (wie etwa die von Wasser und Luft) darzustellen vermögen. Es sei in diesem Zusammenhang auch erwähnt, daß dabei die „frühen“ Laute im Sinne Jakobsons in urtümlicheren Emotionen zur Geltung kommen. Sie finden sich u. a. in den von der Sprachpsychologie erforschten lautlichen Dartellungen gegenständlicher Werte in „primitiven“ Sprachformen. (Hornbostel, Müller, Westermann, Cravens, Kainz.)

7. Über die Ausdruckswirkung eröffnet sich auch eine Brücke von der Phonetik zur Physiognomik, indem sich auf Grund der Werkanalyse Hinweise auf die typologische Ausrichtung der Persönlichkeit ergeben. So wären Petzold und Engelke „typische“ Repräsentanten ergotroper Arbeiterdichtung, Goethe und Eichendorff dagegen Trophotrope. Das wenigste Material zur Ausdruckswirkung, besonders in Hinblick auf die subjektive Lautbedeutsamkeit, findet sich bei den Ideotropen, den Vertretern der Gedankendichtung (Grillparzer, Busch, Morgenstern, Schiller).

8. Die Bedeutsamkeit der Sprachlaute ist ein internationales Phänomen. So gibt es z. B. Hinweise für das Englische (Washburn und Roblee, Eisenson, Souther und Fisher), das Griechische (Sommer) und das Portugiesische (Kayser). Das lebendige Wirken der Lautbedeutsamkeit auch in unserer Zeit zeigt sich in politischen Reden (vgl. Churchills „foul foe“), in Werbeslogans, in emotionalen Ausrufen und in der gefühlsbetonten Alltagsrede. Eine vom Verfasser eben unternommene Erforschung der Internationalität von Gefühlsausbrüchen hat erwiesen, daß etwa im Zorn die Plosive intensiver hervortreten, im Haß dagegen die Frikative, während die Längung der Vokale häufig Wohlbehagen ausdrücken oder erwecken will. Enge Beziehungen zeigen sich zu den Bewegungsspuren des stimmlichen Ausdrucks.

LITERATUR

- Arndorfer, E.: Untersuchungen zur Lautbedeutsamkeit auf Grund von Häufungen gleichartiger Laute, Diss., Wien 1964.
- Cravens, J. M.: Experimentelle Untersuchungen zur sprachlichen Wiedergabe von Sinneseindrücken, Diss., Wien 1956.
- Eisenson J., Souther S. and Fisher J.: The affective values of English speech sounds. In: *The Quarterly Journal of Speech* 26 (1940), S. 580—594 und 29 (1943), S. 457—464.
- Havers, W.: Zur Entstehung eines sogenannten sakralen u-Elements in den indogermanischen Sprachen. In: *Anz. d. öst. Ak. d. Wiss., phil.-hist. Kl.* 84 (1947), S. 139—165.
- Höffe, W. L.: Sprachlicher Ausdrucksgehalt und akustische Struktur. In: *Wiss. Zs. Univ. Jena* 6 (1956/57), Ges.-sprw. Reihe, S. 793—830.
- Hornbostel, E. M.: Laut und Sinn. In: Festschrift für Meinhof, Hamburg (Friedrichsen) 1927, S. 329—348.
- Kainz, F.: Psychologie der Sprache, 2. Band, Vergleichend-genetische Sprachpsychologie, 2. Auflage, Stuttgart (Enke) 1960.
- Kayser, W.: Das sprachliche Kunstwerk, 4. Aufl., Bern (Francke) 1956.
- Müller, H.: Experimentelle Beiträge zur Analyse des Verhältnisses von Laut und Sinn, Diss., Leipzig 1935.
- Schneider, W.: Über Lautbedeutsamkeit. In: *Zs. für dt. Philologie* 63 (1938), S. 138—179.
- Sommer, F.: Lautnachahmung. In: *Idg. Forschungen* LI (1933), S. 229—268.
- Trojan, F.: Der Ausdruck der Sprechstimme, 2. Auflage, Wien und Düsseldorf (Maudrich) 1952.
- Trojan, F.: Sprachrhythmus und vegetatives Nervensystem, Wien und Meisenheim 1951 (Die Sprache, Beiheft 2).
- Trojan, F.: Der Gefühlsresonanztest. In: *Wiener Archiv f. Psychologie, Psychiatrie und Neurologie* V (1955), S. 22—38.
- Trojan, F.: Zeichen, Silbe und Laut in entwicklungsgeschichtlicher Sicht. In: *Phonetica* 1 (1957), S. 63—81.
- Trojan, F.: Zur entwicklungsgeschichtlichen Beziehung des Vokalismus zu den Registern. In: *Aktuelle Probleme der Logopädie und Phoniatrie* I, Basel und New York (Karger) 1960, S. 41—47.
- Washburn, M. F. und Roblee, L.: The affective values of articulate sounds. In: *Am. Journ. Psychol.* 23 (1912), S. 579 ff.
- Werner, H.: Grundfragen der Sprachphysiognomik, Leipzig (Barth) 1932.
- Westermann, D.: Laut und Sinn in einigen westafrikanischen Sprachen. In: *Archiv f. vergl. Phonetik* I (1937), S. 154—172 und 193—222.

РЕЧЕВОЙ АКУСТИЧЕСКИЙ СИГНАЛ И ЕГО ВОСПРИЯТИЕ¹

В. А. АРТЕМОВ*

1. Следует различать акустические свойства и воспринимаемые качества речевого сигнала (2,3). К числу первых относятся число колебаний основного тона, интенсивность и время, а также общая произносительная энергия и ее распределение по полосам частот, перерывы в звучании. Вторые составляют многочисленные воспринимаемые на слух особенности речи, — ее громкость, высота тона и длительность; образы фонем, их вариантов; образы интонаций различных синтаксических и модальных значений; воспринимаемые паузы и расчлененность речи на синтагмы и т. п.

2. Физические свойства речевого сигнала сами по себе, как материальные явления, непосредственно не содержат никаких языковых значений и смыслового содержания, тем более не являются речевыми поступками, т. е. не носят второсигнального характера, который они приобретают в результате научения (1, гл. 4, 5, 6). В этом смысле языковой код подобен любому другому коду, служа лишь переносчиком информации, но не будучи ею сам по себе.

3. Однако языковой код, как и любой код, является не хаотическим собранием разрозненных элементов, а целостной структурой модуляций, организованной (расчлененной, дифференцированной и интегрированной) по системе языка, мысли и речевых поступков. Поэтому электро-акустический и аудитивный анализ речевого сигнала служит индикатором системы языка, мысли и речевых поступков. Применение современных спектрографов и интонографов с использованием счетно-вычислительных машин (4,5) позволяет сделать электро-акустический анализ речи достаточно надежным. Аудитивный анализ весьма субъективен.

4. Речевой акустический сигнал является случайным процессом, он многоинформативен и поэтому чрезвычайно избыточен, что затрудняет его анализ и синтез (6, 7, 8, 9). Речевой акустический сигнал является надежным

* ЛЭФИПР I МГНШИА, Москва.

¹ В русском языке имеются два слова: язык и речь (language and speech, le langage et la parole). Языком мы называем важнейшее средство общения, а речь — процесс общения посредством языка. Язык есть система фонетических, лексических, грамматических и стилистических единиц и правил пользования ими в речи (1, гл. 2).

индикатором: 1. системы языка (10), 2. системы мыслительных категорий (11, 12), 3. смыслового содержания (13), 4. системы речевых поступков (14), 5. особенностей личности его автора (15, 16, 17) (типологических и индивидуальных). (Показ постатейного электроакустического анализа речевого сигнала.)

5. Речевой акустический сигнал является переносчиком информации: 1. о звуковой системе языка (18), 2. о некоторых лексических значениях (19, 20), 3. о синтаксическом строе языка (10, 21) и 4. о стилистических особенностях языка (22).

6. Информация о звуковой системе языка многогранна. Весьма естественно, что математическим путем с применением счетно-вычислительных машин удастся найти акустический коррелят системы фонем и фонемных инвариантов, интоном и интонационных инвариантов (23, 24). Тем самым математически снимается кажущееся недиалектическое противопоставление фонологии и фонетики (26, 27). (Демонстрация экспериментальных данных и графиков.)

7. Информация о некоторых лексических значениях весьма своеобразна. Она относится к различению значений слов, к переносному и идиоматическому значению слова (20), к значению сложносоставных слов (19), фразеологизмов (28). В разных языках здесь наблюдается различная семантическая дифференциация отдельных акустических модуляций языкового кода (31, 32). (Демонстрация примеров из различных языков, включая китайский и вьетнамский.)

8. Особенно дифференцирована информация в речевом сигнале о синтаксической системе языка. Нам удалось на основании электроакустического анализа интонограмм установить неизвестные ранее виды и подвиды повествовательных, вопросительных, побудительных и восклицательных типов предложений, а также решить некоторые весьма спорные вопросы синтаксиса (10), например, вопрос о самостоятельности восклицательного типа предложения (29), о присоединительных конструкциях (13) и прочие, а также, в известной мере, коснуться вопроса о синтаксической модальности (30). (Демонстрация экспериментальных данных на разных языках.)

9. В акустических характеристиках речевого сигнала содержится также информация о стилистических особенностях языка (22). Она относится к характеру предикаций и акустически выраженной экспрессии, как относительной степени отклонения от нейтральных (нулевых) значений частоты основного тона, интенсивности и длительности, а также их уровней, диапазонов и интервалов.

10. Информация о системе сознания содержит данные о смысловых категориях (11, 12), о системе чувств (33, 34) и о системе воли (35).

Речевой сигнал информирует о различных планах мысли (смысловых отношений): существование (экзистенция), наличие (ассерторичность),

категоричность, различные виды проблематичности, причинно-следственные отношения, повествование, вопрос и др. (10, 36).

Действительностью (средством реализации) чувств служит язык, особенно речевая интонация, акустический анализ, который служит надежным индикатором наличия и системы эмоциональных состояний. То же самое относится и к воле. Нам удалось таким именно методом установить систему чувств (34) и зонную природу их распознавания (33). В настоящее время ведется аналогическое исследование воли (35).

11. Смысловое содержание речи реализуется лексически, синтаксически, фонетически и стилистически (37). Адекватное понимание письменного текста или устного сообщения — чрезвычайно сложный процесс вербальной коммуникации. Для этого необходимо не только знание языка, но и всесторонний анализ текста, большого контекста и подтекста, а также реалий (1, гл. 4; 37, 38, 39).

12. Речевой акустический сигнал переносит информацию и о речевом поступке. Это очень сложный вопрос коммуникативной теории и коммуникативного анализа речи (1).

С одной стороны, именно поступки людей определяют языковые характеристики речи. Само высказывание об общественных отношениях в таких условиях становится очень сложным речевым поступком, определяющим звуковые, интонационные, лексические, синтаксические и стилистические особенности речи. В этом смысле в коммуникативной цепи речи ее поступочное звено является решающим для акустических характеристик речевого сигнала и для его восприятия.

В то же время речевой поступок уже как бы выходит за пределы языка. Но ведь то же самое относится и к системе мыслей языкового сигнала.

13. Комплексный (электроакустический, математический, лингвистический, психологический и физиологический анализ) речевого сигнала, проводимый в ЛЭФИПР за последние годы, убедил нас в высокой диагностической ценности этого сигнала в отношении его автора. Во-первых, таким именно путем устанавливаются механизмы и особенности овладения звуковым составом языка человеком, начиная с первого дня его рождения (14, 40, 41). Во-вторых, комплексный анализ речевого акустического сигнала служит диагностическим средством (индикатором) типа высшей нервной деятельности человека (15). В-третьих, такой анализ речевого сигнала позволяет идентифицировать его автора по языковым (42), речевым (43), типологическим и персональным особенностям (46). Это очень сложный математический анализ акустических свойств речевого сигнала (24, 42).

14. Восприятие речевого акустического сигнала — очень сложный психический процесс, его обычно называют теперь аудированием (1, гл. 7; 45). Насколько нам известно, впервые аудиторы были введены в экспери-

ментально-фонетические исследования в 30-х годах в ЛЭФИПР (44). Это было методическое оформление беспорядочного прослушивания изучаемого материала самим исследователем. С того времени накоплено немало надежных закономерностей аудирования как у нас, так и в других лабораториях. Известно следующее. —

15. Аудирование речевых сигналов основывается на его физических свойствах, но не определяется ими однозначно (46).

16. Сами по себе акустические свойства речевого сигнала непосредственно не распознаются. Они распознаются преобразованиями в воспринимаемые качества по системе языка (3). Эта зависимость настолько велика, что одни и те же акустические свойства речевого сигнала воспринимаются по разному лицами, знающими и не знающими данный язык (47). Система языка маскирует дефекты речевого сигнала и надежно противостоит помехам при его распознавании (48, 49).

При прочих равных условиях, высокочастотные речевые сигналы идентифицируются лучше, чем низкочастотные (50, 51). Увеличение длительности также способствует адекватности восприятия речевого сигнала, но в известных условиях идентификация двухсложных слов может достигнуть 94 % при оптимально благоприятном его звуковом составе и артикуляции (50). Наилучшие результаты дает опознавание фраз, состоящих из 8—10 слогов (50).

Резкие скачки распознавания речевого сигнала происходят при переходе с одного языкового уровня на другой — от фонемного уровня к морфологическому, от последнего к лексическому и от лексического к фразовому (50).

Фактор лингвистической частотности не оказывает влияния на распознавание фонем и слогов (50), но он важен при распознавании слов (48).

Речевые сигналы, принадлежащие одному и тому же диктору, идентифицируются лучше, чем принадлежащие различным дикторам (50, 54).

Распознавание речевого сигнала отличается высокой помехо-устойчивостью, которая возрастает по мере языкового и смыслового усложнения речевого сигнала.

Преодоление помех существенно зависит от напряжения произвольного внимания и волевого усиления (49).

17. Восприятие речевого сигнала осуществляется по системе языка (36, 48, 52). Поток речи мы членим на звуки по системе фонем данного языка, не воспринимая тех звуков, которые не соответствуют этой системе. Например, и в русской речи мы произносим одни гласные кратко, а другие длительно, но мы не различаем этого, так как в системе русского языка нет дистрибуции гласных по длительности.

Мы членим также поток речи на слоги по системе морфем данного языка, на слова по системе лексем, на фразы по системе предложений. В отдельных

случаях влияние системы языка на восприятие речи может быть сильнее, чем её акустический или артикуляционный состав (47, 48, 52).

18. Восприятие речевого сигнала обусловлено ситуацией и задачей общения, иначе говоря, речевым поступком (53). Это тесно связано с зонной природой распознавания речевого сигнала (33). Один и тот же акустический речевой сигнал в одной ситуации распознается как просьба, в другой как мольба, а в третьей как дружеский приказ. В одной ситуации слово „дурочка“ имеет социально-отрицательное, а в других — социально-положительное значение.

Коммуникативная задача может сделать восприятие речевого сигнала как угодно избирательным, в частности, аналитическим, синтетическим или смешанным. Задача коммуникации может заставить аудитора считать различительными признаками какие угодно акустические характеристики речевого сигнала (10, 36).

19. Восприятие речевого сигнала существенно обусловлено индивидуально-типологическими особенностями аудитора, в частности, особенностями его слуховых ощущений, восприятий и возможностью хранить воспринимаемый образ в кратковременной памяти, сличая его с эталоном, хранящимся в длительной памяти. Существенна также адекватная задаче исследования установка на аудирование, опыт в аудировании, способность сочетать аналитическое восприятие с синтетическим, знание системы языка, степень практического владения им, умение находить нужные вспомогательные приемы опознавания речевого сигнала (36, 54).

20. Столь сложен и многоинформативен акустический речевой сигнал и столь многогранно его восприятие. Поэтому подлинно научное исследование этой проблемы требует многих творческих усилий коллектива, состоящего из акустиков, психологов, физиологов, лингвистов, математиков и специалистов по теории связи, теории коммуникации и кибернетике, с применением высоко совершенных анализаторов и счетно-вычислительных машин. Одностороннее решение этой комплексной проблемы, особенно, если она покоится на какой-либо случайной прагматической потребности, т. е. носит эмпирический характер, может привести к ложным результатам.

СПИСОК ИСПОЛЬЗОВАННЫХ ИССЛЕДОВАНИЙ И ЛИТЕРАТУРЫ

1. Артемов, В. А.: Психология обучения иностранным языкам, ч. I, гл. 2. Москва, 1966.
2. Артемов, В. А.: Об интонации. Сб. „Вопросы интонации“. Москва, 1953.
3. Артемов, В. А.: О взаимоотношении физических свойств, воспринимаемых качеств, языковых значений и смыслового содержания речи, *ZsPh* 11, 4, 1958.
4. Артемов, В. А.: Метод структурного анализа речевых спектров. Ротапринт I МГПИИЯ. Москва, 1958.

5. Артемов, В. А.: Метод структурного анализа речевой интонации. Ротапринт I. МГПИИЯ. Москва, 1958.
6. Харкевич, А. А.: Очерки общей теории связи. Москва, 1955.
7. Cherry Collin: On Human Communication, London, 1957.
8. Фант, Г.: Акустическая теория речеобразования. Изд-во „Наука“. Москва, 1964.
9. Flanagan, J. L.: Speech Analysis, Synthesis and Perception. Springer-Verlag, New York, 1965.
10. Артемов, В. А.: Экспериментально-фонетические исследования за 50 лет (1917 до 1967). Готовится к печати.
11. Клычникова, З. И.: Информация о мысли в речевой интонации. Тезисы докладов на II-ом съезде общества психологов. Изд-во АПН РСФСР. Москва, 1963.
12. Клычникова, З. И.: К вопросу о понимании речи. Тезисы докладов на научной конференции, посвященной 100-летию со дня выхода в свет руда И. М. Сеченова „Рефлексы головного мозга“. Одесса, 1963.
13. Мильк, В. Ф.: Интонация присоединения в английском языке сравнительно с русским. Канд. дис. Москва—Ленинград, 1957.
14. Тонкова-Ямпольская, Р. В.: В вопросу о становлении речи. Доклад на V-ом Международном конгрессе по фонетическим знаниям. Мюнстер, 16—23 августа 1964.
15. Маланова, Л. Ф.: Интонация речи и некоторые типологические особенности высшей нервной деятельности. Уч. зап. I МГПИИЯ, т. XX. Изд-во МГУ. Москва 1960.
16. Каспарова, М. Г.: Приметные особенности индивидуального произношения согласных. Экспериментальное исследование. ЛЭФИПР, 1965.
17. Ильин, В. И.: К вопросу о вариантах речи. Сб. „Интонация и звуковой состав“. Изд-во МГУ, Москва, 1965.
18. Блохина, Л. П. и Потапова, Р. К.: Ограничение языковых составляющих речи. Экспериментальное исследование, ч. I. ЛЭФИПР, 1965.
19. Степанова, М. Д.: Словоуплощение в современном немецком языке. Ленинград, 1959.
20. Артемов, В. А.: Восприятие и понимание идиом. Уч. зап. I МГПИИ, т. XVIII. Изд-во МГУ. Москва, 1960.
21. Таги-Заде, З. Х.: Сложные предложение в современном азербайджанском языке Автореферат докторской диссертации. Москва—Баку, 1965.
22. Айвазова, Т. А.: Исследование, выполненное в ЛЭФИПР. Москва, 1965.
23. Артемов, В. А.: Об интоне и интонационном инварианте. Сб. „Интонация и звуковой состав“. Изд-во МГУ, Москва, 1965.
24. Татубаев, С.: Статистический способ определения структуры резонансных частот гласных. Сб. „Материалы коллоквиума по экспериментальной фонетике и психологии речи“. Изд-во МГУ, Москва, 1966.
25. Zwirner, E.: Phonometrischer Beitrag zur Geographie der prosodischen Eigenschaften. Proceedings of the 4-th International Congress of Phonetic Sciences, The Hague 1962.
26. Herdan, G.: The Statistics of Phonetic Systems (там же).
27. Трубецкой, Н. С.: Основы фонологии, 1939.
28. Берковская, С. М.: Интонационное оформление пословиц и поговорок в английском языке. Автореферат канд. дис. Москва, 1965.
29. Антипова, А. М.: Интонация восклицания в английском и русском языках. Канд. дис., Москва, 1965.
30. Калиев, Ж. К.: Интонация повествования в современном казахском языке. Канд. дис. Москва, 1967.
31. Румянцев, М. К.: Интонационное выражение завершенности и незавершенности. Сб. „Материалы коллоквиума по экспериментальной фонетике и психологии речи“. Изд-во МГУ. Москва, 1966.
32. Нгуэн Хай Зьонг: Спектральные характеристики гласных и связанные с ними системы тонов во вьетнамском языке. Канд. дис. Москва, 1962.
33. Рожкова, Г. И.: Распознавание эмоциональной интонации. Канд. дис. Москва, 1967.
34. Оксман, Р. А.: К вопросу о классификации чувств. Канд. дис. Москва, 1967.
35. Элиешюте, С. П.: Речевое воздействие на волю. Канд. дис. (ведется исследование).
36. Артемов, В. А.: Психология речи за 50 лет (1917—1967). (Готовится к печати.)
37. Разыграева, Л. И.: Обучение пониманию текста на испанском языке в V классе средней школы. Канд. дис., Москва, 1955.
38. Роговин, М. С.: Проблема понимания. Канд. дис., Москва, 1956.
39. Лучкина, И. М.: Понимание медицинских текстов. Канд. дис. Москва, 1967.
40. Тонкова-Ямпольская, Р. В.: К вопросу о механизмах становления речи. Доклад на IX совещании по проблемам в. п. д. Ленинград, 15—20 апреля 1963.
41. Тонкова-Ямпольская, Р. В.: Исследование процесса становления речи методом электроакустического анализа. Доклад на XVIII Международном конгрессе психологии. Москва, 1966, том III, стр. 95 (тезисы).
42. Блохина, Л. П. и Потапова, Р. К.: Ограничение языковых характеристик речи. Экспериментальное исследование, ч. II, ЛЭФИПР, 1966.
43. Ильина, В. И.: Речевые характеристики коммуникации посредством языка. Экспериментальное исследование, ЛЭФИПР, 1965.
44. Артемов, В. А.: Экспериментальная фонетика и психология в обучении иностранному языку. Сб. того же названия. Изд-во I МГПИИЯ. Москва 1940.
45. Кочкина, З. А.: Об аудировании речи. Психология, 1964, № 3.
46. Артемов, В. А.: Идентификация речевого сигнала человеком. Исследование, ЛЭФИПР, 1966.
47. Селезнева, И. С.: Некоторые особенности восприятия ударения в словах русского языка. „Вопросы психологии“, 1957, № 4.
48. Зимняя, И. А.: К вопросу о воспитании речи. Канд. дис., Москва 1961.
49. Каспарова, М. Г.: Идентификация речевого сигнала в условиях белого шума. Исследование, ЛЭФИПР, 1966.
50. Блохина, Л. П. и Потапова, Р. К.: Идентификация речевого сигнала в зависимости от звуковой структуры и лексического состава. Исследование, ЛЭФИПР, 1966.
51. Зимняя, И. А.: Идентификация речевого сигнала в зависимости от фильтрации верхних и нижних частей спектра. Исследование, ЛЭФИПР, 1966.
52. Каспарова, М. Г.: Восприятие паузы. Автореферат канд. дис. Москва, 1964.
53. Артемов, В. А.: К вопросу об интонации русского языка. Уч. зап. I МГПИИЯ, т. VI. Изд-во МГУ. Москва, 1963.
54. Селезнева, И. С.: Условия, способствующие постоянству произнесения и распознавания речевого сигнала. Исследование, ЛЭФИПР, 1966.
55. Артемов, В. А.: Психология обучения иностранным языкам. Курс лекций. Изд-во МГУ. Москва, 1966.

SUR LE RÔLE DE LA FRÉQUENCE DANS LA PERCEPTION DE L'ACCENT EN ROUMAIN

ANDREI AVRAM*

On a publié, ces dernières années, de nombreux travaux concernant le problème des facteurs physiques qui sont à la base des contrastes accentuels dans les langues possédant un soi-disant accent „d'intensité“.¹ C'est surtout le problème de l'importance relative, dans la perception, de l'intensité, de la durée et de la fréquence qui a attiré l'attention des chercheurs. Sans nous proposer d'établir, pour le roumain, une hiérarchie des trois facteurs énumérés, nous allons présenter quelques remarques sur le rôle de la fréquence dans la perception de l'accent en roumain, après avoir étudié, du même point de vue — dans un article récent² —, le rôle de la durée des voyelles.

Les matériaux que nous avons utilisés ont consisté en 22 phrases (formées, parfois, d'un seul mot), dont un mot (respectivement une syllabe) portait l'accent; les phrases ont été prononcées par un sujet (homme) parlant le roumain littéraire. Les données concernant la durée, la fréquence et l'intensité ont été obtenues à l'aide d'un sonographe, d'un „trans pitch meter“ et d'un „intensity meter“ (les deux derniers appareils ont été connectés à un mingographe).³

Voici quelques-uns des résultats de nos recherches:

1. Dans 14 des 22 cas, le sommet de fréquence coïncide avec la syllabe accentuée. Dans 16 cas, la syllabe accentuée est, à la fois, la plus longue dans la phrase; à noter cependant que dans 2 phrases (*Amar* „amer“ et *Amar?*) la syllabe accentuée (et dont la durée est plus grande que celle de la syllabe inaccentuée) contient trois sons, tandis que l'autre est formée d'un seul son (abstraction faite de ces deux cas, dans les autres phrases toutes les syllabes ont la structure „consonne + voyelle“). Quant

* Centrul de cercetări fonetice și dialectale al Academiei R. S. România, Bucarest.

¹ Voir, par exemple, André Rigault, dans *Proceedings of the Fourth International Congress of Phonetic Sciences...*, La Haye, 1962, pp. 735—748 et les travaux qui y sont cités.

² *Durata vocalelor și perceperea accentului în limba română*, dans „Studii și cercetări lingvistice“ XVII, 1966, 3, pp. 263—269.

³ Les enregistrements ont été faits à l'Institut de phonétique de l'Université de Lund; nous tenons à exprimer notre gratitude à M. Bertil Malmberg, directeur de l'Institut, qui a bien voulu mettre à notre disposition les appareils du laboratoire. Pour des détails concernant les moyens techniques utilisés, voir Kerstin Hadding-Koch och Lennart Petersson, *Instrumentell fonetik. En handledning*, [Lund, 1965], pp. 90—107.

à la coïncidence entre l'accent et le sommet d'intensité, elle ne se produit que dans 8 phrases.

2. Dans tous les cas (8 phrases) où le sommet d'intensité se trouve dans la syllabe accentuée, cette syllabe est, en même temps, caractérisée par un sommet de fréquence. On peut déduire des chiffres mentionnés précédemment (sous 1) que la réciproque n'est pas vraie; par exemple, dans *N-äre mama* „La mère n'(en) a pas“ il y a coïncidence entre accent, maximum de fréquence et maximum d'intensité, mais cette coïncidence ne se produit plus dans *N-äre máma* „C'est la mère qui n'(en) a pas“, où seul le sommet de fréquence se trouve dans la syllabe accentuée [ma], la syllabe la plus intense étant [re].

3. Un fait qui mérite d'être noté est le suivant: 6 des 8 phrases où il n'y a pas coïncidence entre la place de l'accent perçu et le sommet de fréquence sont des phrases interrogatives (dans ces cas le maximum de fréquence se trouve dans la syllabe finale, inaccentuée). Par conséquent — si l'on laisse de côté les phrases où la montée du ton fondamental dans une syllabe inaccentuée répond à la nécessité d'exprimer la signification „interrogation“ — on peut affirmer que l'accent et le maximum de fréquence coïncident dans presque la totalité des cas (le temps ne nous permet pas de discuter ici le problème du rapport entre la place de l'accent et les types d'intonation).

4. Une dernière remarque. On sait que la nature des sons joue, elle aussi, un rôle dans la perception de l'accent.⁴ Il en résulte qu'il est intéressant d'avoir à sa disposition des cas où l'influence de ce facteur soit exclue (comme dans le mot *mama*, par exemple, où les deux syllabes ont une structure identique — évidemment, abstraction faite des traits „prosodiques“). Mentionnons, dans cette ordre d'idées, que parmi les 14 phrases où le maximum d'intensité ne coïncide pas avec la syllabe accentuée il y en a 4 qui présentent un sommet d'intensité dans une syllabe composée de sons identiques à ceux de la syllabe accentuée. Un exemple: dans *Mamaia* (nom d'une localité) le sommet d'intensité se trouve dans la première syllabe, inaccentuée.

Ainsi que les travaux consacrés à d'autres langues nous faisaient le supposer, on constate qu'il n'y a pas, en roumain, de rapport constant et nécessaire entre le maximum d'intensité physique et la place de l'accent perçu. Pour ce qui est des deux autres facteurs, il serait impossible de tirer des faits qui viennent d'être présentés des conclusions concernant l'importance de la fréquence par rapport à celle de la durée (dans 11 des 22 phrases, le sommet de fréquence coïncide avec la syllabe ayant la plus grande durée). Notons cependant que les expériences décrites dans notre article précité semblent prouver que la durée joue, à elle seule, un rôle assez peu important.

Seule la synthèse pourrait nous permettre d'établir, d'une manière précise, l'importance de chaque facteur pris à part et aussi des combinaisons durée + fréquence,

⁴ Cf. D. B. Fry, dans „*Language and Speech*“ I, 1958, 2, pp. 126, 128.

durée + intensité, intensité + fréquence. Ce qu'on peut affirmer dès maintenant c'est que la fréquence (donc le facteur „musical“) joue, dans la perception de l'accent en roumain un rôle considérable et, en tout cas, beaucoup plus grand que ne le laissent supposer les descriptions du roumain dont on dispose jusqu'à présent.

DISCUSSION

Fischer-Jørgensen:

Dans les investigations de cette sorte il faut distinguer entre des types d'accent différents (accent de mot, accent syntaxique, accent emphatique). Ces types n'ont pas seulement des fonctions différentes mais aussi souvent des manifestations phonétiques différentes.

Gsell:

Monsieur Avram a fait une étude très intéressante qui semble montrer qu'en roumain l'accent est réalisé à l'aide de paramètres multiples, avec prédominance de la fréquence. Cette situation a été signalée dans d'autres langues, mais pour être certain, j'aurais voulu:

1. demander à Monsieur Avram, s'il a tenu compte des valeurs spécifiques dont l'introduction modifie considérablement les résultats (g. discussion entre Jassem et Wodarz);
2. lui signaler (une fois prises les précautions) que dans les langues où les 3 paramètres sont assez rapprochés on a intérêt à convertir les valeurs „objectives“ en unités différentielles subjectives. Ce traitement que j'ai élaboré a été appliqué avec succès dans le bel article de H. Rakotofringa, „Sur l'accent distinctif en Malgache merina“, *Langage et Comportement*, vol I, 4, p. 167—171: Valeurs objectives et valeurs significatives.

Rossi:

Dans quel contexte avez-vous enregistré les mots sur lesquels porte votre analyse? Selon vous, la fréquence est la composante essentielle de l'accent. Je crains que vous confondiez les niveaux: accentuel et intonatif. D'autre part vous considérez l'intensité comme une composante mineure. Or vous savez que cette dernière est en fin de compte un artéfact. Il est indispensable de convertir l'intensité objective en sonie en tenant compte de l'intensité spécifique de l'intégration temporelle de l'intensité etc... Vous aboutirez alors à une hiérarchie différente des composantes de l'accent.

Avram:

ad Gsell: Je n'ai pas tenu compte d'une manière systématique des valeurs spécifiques; cependant, dans les phrases étudiées, la majorité des syllabes contenaient la même voyelle [a]. Je remercie M. Gsell d'avoir attiré mon attention sur sa méthode de conversion des valeurs objectives en unités différentielles subjectives (je suis d'accord que cette opération est très importante).

ad Rossi: J'ai étudié, parallèlement, des paires de mots et de phrases dont la syllabe accentuée était la même, mais dont l'intonation était différente, en distinguant ce qui, dans les variations de hauteur est lié, d'une part, à la place de l'accent — de l'autre, à la différence énonciatif — interrogatif. La crainte de M. Rossi n'est pas justifiée; son point de départ reste obscur. Le contexte nous a permis, dans la majorité des cas, de faire des comparaisons entre des syllabes accentuées et des syllabes inaccentuées ayant une structure similaire, sinon identique (*Máma. Mamáia? Máma ia?* etc.); on a donc éliminé, dans la mesure du possible, l'intervention de l'intensité spécifique (évidemment, le problème devrait être étudié plus en détail, en utilisant des moyens que je n'ai pas eu à ma disposition). Quant à la hiérarchie des trois facteurs, M. Rossi m'attribue à tort l'affirmation suivant laquelle „la fréquence est la composante essentielle de l'accent“; je n'établi aucune hiérarchie (et je l'ai dit expressément dans ma communication).

К ХАРАКТЕРИСТИКЕ ИНТОНАЦИИ ПОВЕСТВОВАТЕЛЬНОГО ПРЕДЛОЖЕНИЯ В СЛАВЯНСКИХ ЯЗЫКАХ

А. И. БАГМУТ*

Задачей сравнительного изучения интонации славянских языков является, прежде всего, исследование конкретных вопросов интонационного построения фразы в зависимости от типа предложения и его синтаксического строя. В докладе на материале простого повествовательного предложения украинского и чешского языков, взятых в ситуационном контексте, прослеживаются основные интонационные черты этих языков.

Можно считать установленным (что подтверждают и наши наблюдения), что в повествовательных предложениях русского, украинского и чешского языков, произнесенных без особой эмоционально-волевой окраски и специального логического выделения какого-либо центра высказывания, фразовое ударение находится в конце предложения (Л. В. Щерба, И. П. Раппопов, Фр. Данеш). Такое „автоматизированное“ фразовое ударение характерно не только для элементарных предложений типа укр. „*Насували хмари*“, „*Зосталася Маруся сама*“ или чеш. „*Někdo zvoní*“, „*Jirka je ještě malý kluk*“, но и для распространенных предложений („Робочий день Миколи Ласточкина починався із сходом сонця“, „Стрілянина в самісінькому центрі ворожого розташування породила папіку на цілій ділянці“, „*Srdce jí znovu sevřel pocit úzkosti*“, „*Za pět minut budou hodiny odbíjet desátou*“ и под.).

Фразовое ударение в этих предложениях создает, таким образом, некоторое единообразие в интонационном оформлении повествования как распространенных, так и однословных предложений.

При экспериментальном исследовании интонации постоянное ударение конца повествовательной фразы выражается увеличенными темпоральными характеристиками и довольно значительным для конца фразы изменением интенсивности произнесения ударного гласного (3—4 мм). Движение основного тона (мелодия предложения) остается в большинстве случаев нисходящим, с отрицательным интервалом (напр. 80—70 гц, 110—80 гц — мужской голос). В более редких случаях движение основного тона ударного

* Отдел экспериментальной фонетики Института языковедения АН УССР.

гласного имеет восходяще-нисходящее направление (напр., 70—80—70 гц).

„Автоматизированное“ ударение, воспринимаемое на слух довольно четко, вследствие своего положения в конце фразы выделяется по интенсивности произнесения и высоте основного тона сравнительно небольшими абсолютными величинами, которые в речи, однако, получают значение основных смысловых различительных дифференцирующих признаков.

Интонационные явления конца фразы можно противопоставить явлениям начала фразы, выражающимся в абсолютно больших величинах и соответственно имеющим меньшее дифференцирующее значение. Для начала фразы характерна наиболее яркая картина распределения основных физических характеристик интонации (в том числе и более быстрый темп, что особенно заметно на материале украинского языка и менее характерно для чешского языка). Начало фразы, обычно более сжатое темпорально, произносится с наибольшей в предложении интенсивностью и со значительными восходящими интервалами движения основного тона (напр. 80—170 гц 100—200 гц). Такие яркие характеристики начала повествовательного предложения по сравнению с умеренными характеристиками середины и конца предложения проявляются в речи независимо от длины фразы и от того, является ли ударный слог начальным во фразе или же ему предшествуют предупредительные слоги.

Рассматривая явления начала фразы в украинском и чешском языках в сравнительном плане, следует отметить, что постоянное ударение в чешском языке еще как бы усиливается основными физическими характеристиками, в результате чего движение основного тона получает более резкое восходящее направление, тогда как в украинском языке движение основного тона в таких предложениях более сглажено, более плавное, вплоть до появления иногда почти равного тона во всей фразе. Изменения в интенсивности произнесения и в чешском и в украинском языке выражены довольно ярко и имеют общую тенденцию постепенного ослабления интенсивности произнесения к концу фразы. Отметим, что и в украинском и в чешском языках наблюдается некоторый подъем основного тона на заударном слоге, более яркий в начале фразы, сопутствуемый увеличением интенсивности.

Таким образом, общий интонационный контур повествовательных предложений — с ярко выраженными показателями физических характеристик начала фразы, имеющими относительно небольшое собственно языковое значение, и со значительно меньшей в абсолютных величинах (кроме длительности звучания, которая возрастает), но дискретно очень показательной характеристикой конца фразы — является типичным для исследуемых языков. Конкретное его проявление варьируется в связи со своеобразием языков — наличием в них постоянного или непостоянного ударения, своеобразием в интенсивности произнесения, характерной для данного

языка, темпоральными особенностями, связанными, в частности, с выделением ударных слогов, с характером словесного ударения и др.

DISCUSSION

Аванесов:

Настало время на изучение интонации распространить лингвогеографический метод. Интонация в пределах одного языка и чем более группы языков нередко отличается. Русское повествовательное предложение существенно отличается в интонационном отношении по диалектам. Отличается интонация вопроса в русском, с одной стороны, и белорусском и польском. Эти отличия интересны как в структурном отношении, так и в историческом (возможно субстратные явления).

Mahnken:

(in russischer Sprache): Hier wurden nur zwei Beispiele einfacher Aussagesätze miteinander verglichen. Um zu einem fundierten Vergleichsurteil über die Intonationstypen zweier oder mehrerer Sprachen zu gelangen, müßte m. E. bedacht werden, daß weitere Varianten des Aussagesatzes bestehen. So hängt z. B. die relative Lage der Akzentsilbe zum vorausgehenden Intonationsverlauf in der terminativen Kadenz im Serbokroatischen u. a. davon ab, wie stark die relative Hervorhebung dieses betreffenden Wortes ist.

ИНСТРУМЕНТАЛЬНО-ФОНЕТИЧЕСКИЕ ДАНИЕ ОТНОСИТЕЛЬНО ИНТОНАЦИИ И СЛОВЕСНОГО УДАРЕНИЯ В ТЮРКСКИХ ЯЗЫКАХ

УЗБЕК БАЙЧУРА

Интонация и ударение в тюркских языках почти не исследовались. Наши данные об ударении и интонации основаны на кимографических (ошибка записи до 2,8 %) и осциллографических записях до 40 дикторов — посетителей 12 тюркских паречий (из всех 4-х групп тюркских языков), сделанных в 1952—1959 гг. в Ленинграде, Казани и др.

1. Общая картина интонации предложения, фразы в тюркских языках сводится к восходяще-нисходящему движению тона — в сложных и простых предложениях и в сочетаниях вопросительного предложения вместе с ответом на него: в первом (вопросительном) тон повышается, во втором (ответном) тон падающий. Ср. уйгурское *išläsäñ tišlärsän* 'Будешь работать — будешь есть' (2 изм.) тон в первом слоге — нуль (гласный выпал), во втором 158 гц., в третьем 194, в четвертом нуль, в пятом 144, в шестом 138 гц и т. д. Однако в этом восходяще-нисходящем рисунке движения тона вершина тона может смешаться к началу или к концу предложения, фразы — вплоть до первого или последнего слога — в зависимости от типа предложения, семантических и логических факторов, места словесного ударения и пр.

2. При этом в простых повествовательных и утвердительных предложениях, ответах на вопросы и во второй (главной) части сложно-подчиненного предложения обычно имеем нисходящий тон, причем максимум высоты тона, как правило, на втором слоге первого двусложного слова, если это слово имеет ударение на конце, а если первое слово односложное или двусложное, но с ударением на первом слоге, — тогда максимум высоты тона будет на первом слоге предложения. Напр., казан-татар. *aš pəšlə* 'Суп сварился' движение тона по слогам будет: 147—143—125 гц., *min qazanda tərat* 'Я живу в Казани' движение тона по слогам: 151—135—129—139—120—92 гц (3). Здесь семантически важная часть *min qazanda* 'Я в Казани' имеет первостепенное тоническое ударение на первом слоге, второстепенное — на последнем слоге этой фразы (-da), совпадающее со словесным ударением, тогда как семантически менее важный последующий глагол *tərat* 'живу' произносится с падающей интонацией, но тесситуре находящейся на более низком уровне.

3. В придаточных предложениях и деепричастных оборотах, предшествующих главному, в вопросительных предложениях вообще (напр. в уйгурском, мишар-татарском) или только при общих вопросах (казан-татарский) тон является восходящим, максимальная высота тона приходится на последний или предпоследний гласный в предложении — в зависимости от места словесного ударения. Напр., каз-тат. *aşpəştə mə* 'суп сварился?' (3) движение тона по слогам: 114—116—134—84 гц. и др. При специальных вопросах в каз-тат. тон нисходящий, ср., напр., *närsä pəstə* 'Что сварилось?': 175—152—117—115 гц. и др., а если вопросительному слову предшествует другое слово, то в последнем тон ниже, чем в первом, напр. каз-тат. *sin qaida tuoansəN* 'Ты где родился?' (2) движение тона по слогам: 152—183—139—120—114—112 гц. Восходящий тон может иногда иметь и ответ на вопрос в татарском.

4. Словесное ударение может оказывать влияние на интонацию предложения. Напр., если в начале предложения идет слово с ударением на первом слоге, — тон начала предложения нисходящий, если же первое слово имеет ударение на втором слоге, — тон начала фразы восходящий, ср. движение тона уйгур. *qandaŋ aş pişti* 'Какой суп сварился?' (3): 158—150—165—0—168 гц., *nimä oçuq* 'Что открыто?' (4): 148—161—176—148 гц. Аналогично в татарском и др.

5. Влияние интонации на словесное ударение отмечено нами лишь в конечном слове предложения: при широком гласном в первом слоге и узком во втором возрастает относительная высота тона первого гласного конечного слова, но сохраняются конечные силовое и экспираторное ударения (напр., уйг., мишар-тат.) и лишь изредка возможен перенос на первый слог слова и тонического и силового ударений, а с другой стороны, возможно и полное выпадение обоих гласных последнего слова в предложении. Напр., в слове *oçuq* 'открытый' (см. выше), *bolsin* 'пусть будет': *az bolsima saz bolsin* 'Пусть будет мало, но хорошо' (4): 158—149—154—154—152—142—134 гц. В казан-татарском, где словесное ударение более устойчиво, мы не зарегистрировали подобного влияния интонации на ударение.

6. Отсутствие инструментальных исследований и презумпция о единообразии ударения в тюркских языках привели к разногласиям среди языковедов, ибо, как установлено нами на основе экспериментальных данных, словесное ударение в тюркских языках является неодинаковым как по месту, так и по характеру. Оно может состоять из многих компонентов (сила звука, высота тона, экспирация и др.) и по отдельным компонентам будет передвигаться с последнего слога слова на первый слог — по мере нашего передвижения от более восточных языков к языкам географически более западным. В уйгурском, казан-татарском, „уфимско-башкирском“, низовом чувашском, а также в бурят-монгольском силовое и тоническое ударения обычно на последнем слоге, может возрасть и квантитет

в нем. В кипчакских (казахский, частично-мишар-татарский) и в верхнем чувашском обычно наблюдается четко выраженное начальное силовое и тоническое ударения, но возрастание квантитета обычно на конечном слоге. Промежуточное положение занимают языки огузские (азербайджанский, туркменский, турецкий), а также тувинский, хакасский, горно-башкирский и др.: в них ударение может падать и на первый и на последний слоги, а большей частью компоненты ударения распределяются между двумя слогами в различной пропорции. Отсутствие четко выраженного словесного ударения приводит в соответствующих диалектах к возрастанию зависимости компонентов ударения от фонетических факторов.

ÜBER DIE ANWENDBARKEIT EINFACHER PROSODISCHER REGELN FÜR DIE SPRACHSYNTHESE

R. BAKIS, E. H. ROTHAUER, D. MAIWALD*

Strategien für die Erzeugung synthetischer Sprache müssen sich mit drei Problemen auseinandersetzen: a) Wahl der Elemente (Lautelemente oder -gruppen, Wörter, Wortgruppen), b) Verkettung von und Interpolation zwischen den Elementen. c) Berücksichtigung der prosodischen Suprastruktur. Zahlreiche Arbeiten haben bereits vielversprechende Teillösungen gebracht, die aber nur die beiden ersten Problemgebiete berücksichtigen und deshalb prinzipiell nur zu Sprache mit begrenzter Qualität führen können. Grundlegende Verbesserungen der Qualität erfordern eine quantitative Erfassung der prosodischen Strukturen, die in der Literatur noch nicht mit der für uns benötigten Genauigkeit beschrieben sind. In qualitativer Hinsicht wissen wir, daß die prosodischen Merkmale der Sprache vor allem durch den Tonhöhenverlauf und die zeitliche Struktur der Elemente gegeben sind. Wir suchen als Arbeitsmodelle einfache Algorithmen zur Bestimmung des zeitlichen Verlaufs der Parameter Tonhöhe und Dauer, um einen vorgegebenen Sprechausdruck zu erhalten. Durch spätere schrittweise Verfeinerung dieser Modelle, vielleicht unter Verwendung zusätzlicher Parameter, sollte ihre Verknüpfung zunächst mit der grammatischen Struktur eines Satzes möglich werden und auf einer anderen, vielleicht höheren Stufe, mit seinem semantischen Gehalt oder mit dem Emotionszustand des Sprechers.

Bei den genannten Untersuchungen verwenden wir für die Sprachsynthese einen Computer-gesteuerten Vocoder.¹ Dadurch haben wir die Möglichkeit, bei einem gegebenen Wort oder Satz die Tonhöhe ohne Beeinflussung der Dauer zu verändern und umgekehrt. In der primitivsten Stufe einer Synthese genügt es demnach, die prosodischen Merkmale Tonhöhe und Dauer für das jeweilige Sprachelement direkt vorzugeben. Hierzu ein einfaches Beispiel, bei dem als Sprachelement das Wort „ai“ verwendet wurde. Bei wirklicher Sprache dürfen die Übergänge zwischen verschiedenen Tonhöhen jedoch nicht so abrupt gemacht werden, außerdem ist es bei der Sprachsynthese nicht sinnvoll, die Vielfalt aller möglichen Tonhöhen zur Verfügung zu stellen.

* IBM Forschungslaboratorium, 8803 Rüschlikon, Schweiz.

¹ E. H. Rothauer: The integrated vocoder and its application in computer systems. *IBM Journal* 10 (1966), 455—461.

Der Verlauf der Tonhöhe kann nach Öhman und Lindqvist²⁾ bei normaler Sprache durch die Kombination dreier Komponenten angegeben werden, nämlich durch die Satzmelodie, eine Folge von Einzelwortmelodien und eine physiologisch bedingte Welligkeit (interaction ripple). Der Verlauf der Satzmelodie kann durch eine zeitliche Folge von nur wenigen, geglätteten Stufenfunktionen angenähert werden. Die Melodieempfindung wird nach Hockett³⁾ durch die Verwendung von vier Tonhöhen-niveaus und den drei Endungen „eben“, „fallend“ und „steigend“ hinreichend gut nachgebildet. Um annehmbar klingende Sprache zu erzeugen, muß man auch ihre zeitliche Struktur beeinflussen. Einige Elemente müssen gedehnt, andere aber in ihrer Zeitdauer verkürzt werden. Wir haben gefunden, daß auch bei sehr unter-

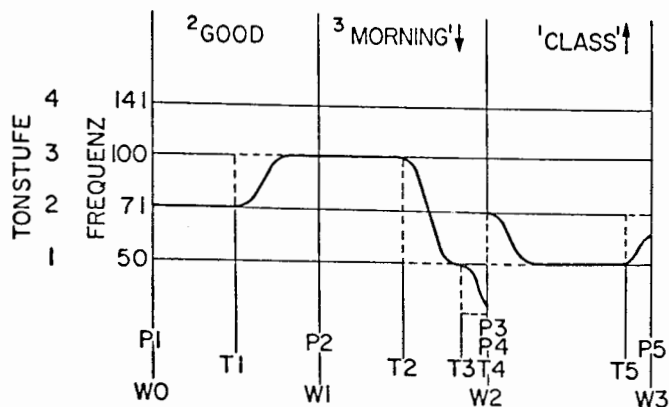


Fig. 1.

schiedlich dauernden Elementen keine Glättung der Geschwindigkeitssprünge erforderlich ist. Nach Hockett³⁾ ist die englische Sprache dadurch charakterisiert, daß bei gegebener Sprechgeschwindigkeit es ebenso lang dauert, um von einer betonten Silbe zur nächsten zu gelangen, ob keine oder viele Silben dazwischenliegen. Es wurde ein Programm entworfen, mit dem im Computer gespeicherte Einzelwörter zu Sätzen zusammengestellt und mit prosodischen Merkmalen versehen werden können, die wir in Anlehnung an die erwähnten Ergebnisse obiger Autoren formuliert haben. Durch die Verwendung ganzer Wörter als Sprachelemente können wir die Einzelwortmelodie und die oben beschriebene Welligkeit als gegeben voraussetzen. Fig. 1 zeigt am Beispiel des einfachen Satzes „Good morning class“ die von uns verwendeten Notierungen und Algorithmen. Die oberste Zeile gibt den gewählten Satz wieder, wie er dem Rechner zugeführt wird, mit den Zahlen für das gewünschte Tonhöheniveau in Halboktavschritten und den Symbolen für die Endungen. Zu

Beginn eines Satzes und auch am Anfang jedes Makrosegmentes existiert ein bestimmtes Tonhöheniveau, für das wir Nummer 2 gewählt haben.

Der Übergang von einem Niveau zum anderen wird geglättet und erfolgt jeweils in der Mitte zwischen den Punkten, bei denen eine Tonhöhenzahl angegeben ist (T_1, T_2 in Fig. 1). Die Übergänge T_3 und T_5 rühren von den Angaben über die gewünschte Endung her. Die Stufe beginnt immer 100 msec vor dem Zeitpunkt, für den die Endung angeschrieben ist. Der Algorithmus für Betonungen lautet: Die Sprechdauer zwischen zwei aufeinanderfolgenden Betonungszeichen wird auf einen konstanten Wert 0,8 sec transformiert. Zuvor wird die Zeitdauer für den unbetonten Teil um den Faktor 2 verkürzt.

Das unmittelbare Ergebnis der vorliegenden Arbeit ist die Erkenntnis, daß einfache prosodische Modelle zwar für eine erste, phänomenologische Beschreibung der Sprache ausreichen mögen, daß aber für eine Synthese von Sprache höchster Qualität wesentlich kompliziertere Modelle erforderlich sind. Wir hoffen, außerdem gezeigt zu haben, daß Arbeiten auf dem Gebiet der Sprachsynthese in ein Grenzgebiet fallen, das für Phonetik, Linguistik, Nachrichtentechnik und Datenverarbeitung gleichermaßen von Interesse ist.

DISCUSSION

Paulus:

Für die praktische Anwendung ist die vorgeschlagene Methode wegen ihrer Allgemeinheit unvorteilhaft. Die Erstellung der Daten erfordert vom Benutzer gründliche Kenntnisse über prosodische Merkmale. Besteht die Absicht, die Benützung durch Verwendung von geeigneten „Prosodemen“ (deren Anzahl überschaubar gering sein muß) zu erleichtern?

Maiwald:

ad Paulus:

Zum ersten Teil der Frage:

Um die Vielfalt der erzeugbaren prosodischen Muster für den Anwender einzuschränken und damit die Programmierarbeit für ihn zu vereinfachen, verwenden wir eine übergeordnete Notierung, wobei — vergleichbar dem Zusammenhang zwischen FORTRAN und Maschinensprache — durch einfache Symbole die nötigen Unterprogramme aufgerufen und durchgeführt werden.

Zum zweiten Teil:

Im Augenblick sind wir dabei, die grammatische Struktur der Sätze einzubeziehen, derart, daß für gleiche grammatische Strukturen gleiche prosodische Muster verwendet werden, unabhängig von den im Satz auftretenden Wörtern. Es ist leicht einzusehen, daß dieses Verfahren zu einer radikalen Einschränkung der Programmierarbeit für den Anwender führen kann. Die Ergebnisse sind recht ermutigend.

² S. Öhman und J. Lindqvist: Analysis-by-synthesis of prosodic pitch contours. *STL-QPSR-4* 1965 Royal Institute of Technology, Stockholm, 1—6.

³ C. F. Hockett: A course in modern linguistics. The Macmillan Company, New York 1958.

QUELQUES OBSERVATIONS SUR LE CONSONANTISME DE LA MODALITÉ CUBAINE DE L'ESPAGNOL

LUBOMÍR BARTOŠ

Dans les dernières années, plusieurs linguistes et dialectologues, citons parmi eux José Pedro Rona, préconisent une nouvelle dialectologie hispanoaméricaine qui partirait des modernes postulats méthodologiques. Dans leur conception de recherches, ils prêtent une grande attention aux faits phonétiques et phonologiques comme base pour l'étude des autres plans de la langue.¹ Il faut se féliciter d'une telle entreprise mais pourtant, à notre avis, l'étude de ces faits ne vaut rien si elle n'est pas réalisée à l'aide des méthodes expérimentales les plus modernes.

Les recherches actuelles dans le domaine de l'espagnol en Amérique démontrent que, en ce qui concerne son aspect phonétique, la situation est beaucoup plus complexe et compliquée qu'on ne le croyait. Ne disposant que de dix minutes pour notre rapport, il nous faut laisser de côté les problèmes théoriques généraux qui se posent dès qu'on aborde l'étude du phonétisme hispanoaméricain.

Quant à la modalité cubaine, comparée à d'autres parlers de l'espagnol en Amérique, elle présente, au point de vue phonétique, un caractère spécifique dû, en premier lieu, à l'importance de l'adstrat nègre. Il faut souligner que les nègres qui sont venus involontairement à Cuba, ont joué un rôle beaucoup plus important que les indigènes dans les autres pays américains.² Il est difficile, bien sûr, de déterminer quels phénomènes phoniques actuels sont le résultat de l'évolution autochtone de l'espagnol et dont il faudrait chercher l'origine dans l'adstrat nègre.³ Mais il est presque hors de doute que dans les riches modulations de la chaîne parlée dans l'espagnol cubain, on peut constater l'influence des langues africaines. Ceci nous amène à nous demander si même dans les habitudes articulatoires des cubains on ne pourrait pas trouver des traces du phonétisme africain.

Dans notre communication nous allons nous limiter à présenter quelques remarques relatives à certains aspects du consonantisme de l'espagnol cubain. Nos conclusions sont fondées sur l'analyse de deux radiofilms enregistrant la prononciation

¹ Voir J. P. Rona, *Aspectos metodológicos de la dialectología hispanoamericana*, Montevideo 1958, p. 25.

² Cf. C. T. Alzola, *Habla popular cubana*, Universidad de La Habana, 159, La Habana 1963, p. 98.

³ Amado Alonso, *Estudios lingüísticos*, Temas hispanoamericanos², Madrid 1961, pgs. 266—267.

de deux sujets cubains appartenant à la classe instruite. L'analyse des radiofilms confirme d'une part quelques observations faites par l'oreille et d'autre part elle indique d'intéressantes modifications de positions et localisations articulatoires. Ces modifications, quoiqu'aujourd'hui sans effet au point de vue auditif, signalent les tendances qui pourront aboutir, dans le futur, à une transformation du système articulatoire.

Il est bien connu que la prononciation des cubains instruits est plus ou moins éloignée de la norme orthoépique représentée par les travaux de T. Navarro Tomás et de Alarcos Llorach. En premier lieu, dans la modalité cubaine ne se réalise pas la fricative interdentale [θ] ce qui a amené la défonologisation de l'opposition entre la fricative interdentale [θ] et la sifflante [s], le soit disant „seseo“, phénomène très répandu dans l'Amérique hispanophone. Il existe cependant une divergence d'opinion en ce qui concerne la réalisation phonétique du produit de cette défonologisation. Selon Canfield⁴ qui tâche de délimiter la distribution territoriale des différentes réalisations de la sifflante, donc dans la plupart des pays de l'Amérique ainsi qu'à Cuba, la sifflante s'articule comme dorso-alvéolaire. Mais nos films démontrent — contrairement à l'opinion de Canfield — que la sifflante est prononcée avec la pointe de la langue s'appuyant contre les alvéoles ou contre la partie postérieure des incisives supérieures, la forme du dos de la langue étant légèrement concave. L'articulation de la sifflante cubaine est donc apico-dento-alvéolaire, aussi supposée la même pour les Antilles par Navarro Tomas.⁵

C'est aussi la consonne latérale dorso-palatale représentée par le graphème <ll> qui est absente dans le phonétisme cubain. La perte de la latérale palatale et sa substitution par la fricative palatale, assez répandues en Amérique, surtout dans la zone antillaise et de La Plata, sont dues à une articulation affaiblie. Puisqu'on suppose cette évolution autonome pour les zones différentes, il n'est pas exclu que dans la zone caraïbe, c'est le phonétisme africain qui a contribué à ce changement.⁶ Nos radiofilms confirment qu'il n'y a aucune différence entre la prononciation des consonnes représentées par les graphèmes <ll> et <y>.

Dans l'espagnol cubain la réalisation vélaire du graphème <j> n'existe pas non plus. Elle est remplacée par la prononciation laryngale, résultat aussi peut-être de la prononciation relâchée. Remarquons cependant que l'on entend la réalisation vélaire dans les noms propres d'origine germanique (Hart, Hollywood).

La réalisation du phonème (b) suscite aussi quelques réflexions. Est-ce que la distribution complémentaire des variantes occlusive et fricative représentant ce phonème est identique dans l'espagnol cubain à celle donnée pour l'espagnol européen? La réponse est négative. Il résulte de nos films que la variante occlusive ne se limite pas à la position initiale et après les nasales, mais qu'on la rencontre aussi dans d'autres

positions, même dans l'intervocalique. Dans ce dernier cas, l'occlusion labiale n'est pas si forte et elle change selon les voyelles voisines.

Les radiofilms ont aussi prouvé notre hypothèse que dans la modalité cubaine existe la prononciation labiodentale pour le graphème <v>. Nous l'avons enregistrée surtout dans les positions initiale et intervocalique (vaya, vaga, ovalo). Il paraît que cette réalisation est assez répandue à Cuba et qu'on ne peut pas la considérer comme une prononciation de maître d'école.⁷

Quant aux consonnes dentales, il est intéressant de noter, d'après notre matériel, une grande variation articulatoire pour la variante fricative [d], constatée déjà par Navarro Tomás. Mais à la différence des croquis de cet auteur où le dos de la langue est convexe, le dos de la langue dans les articulations cubaines est nettement concave. Les [d] et [t] occlusives, dont le lieu de localisation se rapproche des alvéoles, manifestent la même tendance qui nous paraît très caractéristique pour la modalité cubaine.

Une articulation spécifique cubaine constitue le [n] vélaire qui, à la fin des mots, nasalise la voyelle précédente, avec une tendance à la disparition de la consonne. D'ailleurs, on note à l'audition que la modalité cubaine se caractérise par un timbre nasal de la voix.

Il y a d'autres particularités dans le consonantisme cubain qui mériteraient d'être examinées mais le temps de notre rapport étant limité, nous devons nous borner à ce résumé très succinct que nous venons de présenter.

DISCUSSION

Valdman:

Avant d'invoquer les influences d'adstrat dans les traits du consonantisme cubain, il faudrait d'abord étudier les facteurs internes, dans le cas présent, des différences dialectales originelles. Nous avons nous-mêmes constaté que de soi-disantes influences africaines dans le créole français antillais se retrouvent dans les dialectes français de la Louisiane ou du Canada français.

⁴ D. L. Canfield, *La pronunciación del español en América*, Bogotá 1962, pgs. 79—80.

⁵ T. Navarro Tomás, *Manual de pronunciación española*, Madrid 1953, p. 107.

⁶ Cf. aussi Amado Alonso, o. c., p. 210.

⁷ B. Malmberg, *L'espagnol dans le Nouveau Monde — problème de linguistique générale*, *Studia Linguistica* I, II, Lund 1947—1948, pg. 58.

ФРАЗОВОЕ УДАРЕНИЕ, КАК ВОСПРИНИМАЕМЫЙ И АКУСТИЧЕСКИЙ КОМПОНЕНТ ИНТОНАЦИИ

К. К. БАРЫШНИКОВА*

1. В комплексе характеристик звуковой структуры речевых единиц ударение является одним из наименее изученных, несмотря на многочисленные общелингвистические, грамматические и фонетические исследования, проведенные на материале различных языков. Как известно, акцентная структура слов на уровне морфологии уже давно имеет свои ставшие традиционными характеристики для различных языков. Однако природа воспринимаемых качеств и акустических свойств ударения в слове до сих пор остается нераскрытой.

Еще менее определена акцентная структура более крупных речевых единиц на уровне синтаксиса, то есть акцентуация фразы.

2. Структурные исследования интонации фраз — предложений, проведенные в ЛЭФИПР, выявили на многочисленном материале различных языков (немецкого, английского, французского, русского, украинского, грузинского, казахского, узбекского и др.), что выделенность сегментов, воспринимаемая как ударение реализуется на акустическом уровне изменениями одного, двух или трех параметров: частоты основного тона, интенсивности, времени в их различном сочетании.

Одним из постоянных признаков, способствующих идентификации ударения и его степени, как правило, является длительность выделяемого сегмента.

3. Сопоставление данных акустического, перцептивного и лингвистического анализа дало возможность выделить предикативную, делимитативную, экспрессивную и ритмическую функции ударения как компонента интонационной структуры фразы. При этом каждая из этих функций ударения реализуется особыми акустическими признаками. Нашими исследованиями определены предикативная, делимитативная, экспрессивная и ритмическая функции французского языка, указаны их акустические свойства и воспринимаемые качества.

Различение указанных функций ударения возможно только в результате

* ЛЭФИПР I МГНИИЯ им. Мориса Тореза, Москва.

раскрытия специфики взаимодействия словесного и фразового ударения. Взаимодействие ударных и неударных сегментов фразы отражает в сложной структуре фразы специфику морфолого-синтаксического строя данного языка, его фонетических тенденций.

4. Языковые интонационные единицы — интонымы или просодемы включают в свою структуру комплекс характеристик, идентифицируемых на слух, как ударения и их варианты.

5. Дальнейшее изучение ударения, как компонента интонации фразы в контексте требует строгого определения уровня и метода для каждого вида исследования. Необходимо включить в круг вопросов, связанных с изучением ударения, исследования артикуляции фразы, как интонационной акцентной единицы.

SPEECH ERRORS ON DELAYED AUDITORY FEEDBACK

J. BAŠTECKÝ*, O. VINAŘ*, B. BOROVIČKOVÁ, VL. MALÁČ, M. ZICHOVÁ

The study of speech functions can be considered as one of the best methods of investigation of psychic functions. The possibility to evoke a disturbance of a certain function under controlled conditions proved to be a good way to better knowledge of such a function. Therefore we have applied the delayed auditory feedback (DAF, Lee 1950) in experimental psychiatry in a similar way as Goldfarb et al. (1958), Morávek (1961), Spear and Bird (1963) and Hughes et al. (1963).

Our first hypothesis was that the disturbance induced by delayed auditory feedback (DAF) would be smaller in persons who are more independent on exteroceptive signalisation. The intensity of the disturbance of speech during DAF could be a measure of withdrawal. It is very difficult to find a sufficiently valid and reliable criterion of the deepness of withdrawal although it is one of the classical psychiatric syndromes.

Therefore in the beginning of our work we decided to compare the effect of DAF in 3 groups of persons:

1. healthy experimental persons
2. these same persons in LSD (lysergic acid diethylamide) induced state
3. schizophrenic patients

Our hypothesis was the following: the disturbance of speech induced by DAF would be the greatest in healthy persons, smaller in these persons after ingestion of 100 γ of LSD and the smallest in schizophrenics.

This hypothesis was not confirmed, but some other results of our work (Vinař et al. 1965a, Vinař et al. 1965b, Baštecký 1965, Baštecký et al. 1966) show that the method of DAF can contribute to the differentiation of groups of psychiatric patients from healthy controls, that it reflects the effects of some psychotropic drugs and that it can be used as a predictive tool of success of pharmacotherapy in psychiatry.

In the present paper we would like to analyse the kind of speech errors induced by DAF and look for differences among the groups of examined persons.

* Psychiatric Research Institute, Prague 8-Bohnice.

Method: The examined persons read a selected text, consisting of 7 parts and repeated the reading twice. During the reading the delay is presented alternately so that every part is read under normal conditions and with delay. The Jansen tape recorder is used and a delay of 142 msec. and intensity of 65 db are applied. Beside the speech errors, following parameters are registered: speed of reading, vocal intensity measured in db and electrical activity of mimic muscles.

The speech errors are registered by two independent raters listening to the tape record. The following kinds of errors are recorded: repetitions, exchanges, tonic spasms, additions and omissions which are considered as errors in the region of articulation (errors of the type A) and acceleration, deceleration and changes in intonation which are considered as errors in the region of prosodic features (errors of the type B). The inter-rater reliability of this procedure is sufficiently high, and in rare cases of disagreement, there are taken into consideration the mean numbers of recorded speech errors.

Subjects: three groups of persons were examined:

1. 65 healthy controls, 40 women and 25 men, medical students and students of psychology, nurses and physicians, all without any marked mental disorder. The mean age of this group is 25,7 years.

2. 135 schizophrenic in-patients of the Psychiatric Research Institute in Prague, 81 women and 54 men. The diagnosis of schizophrenia had to be confirmed unanimously by all—at least three—psychiatrists on a case conference and by Fould's RSSSI (1963). The examination was performed in the first fortnight after admission of the patient before treatment, i.e. after at least 10 days without any active medication. The mean age of the patients is 32,8 years.

3. 37 persons in a LSD-induced state three hours after ingestion of 100 μ of lysergamide. This group was selected from the control group, they were 18 men and 19 women. Their motivation for the inclusion in this experimental group was a professional interest in LSD experience. The mean age of this group is 22,7 years.

The difference in the age among the compared groups is not statistically significant.

Results: The increase of the absolute number of all speech errors induced by DAF does not differ when we compare the controls with the persons in a LSD induced state, the mean increase is about the same: 27,41 in controls and 27,86 in LSD induced state. However, in the group of schizophrenic patients this increase is greater ($P < .01$) (the mean difference being 38,97) than in both previous groups.

When we separate the two types of errors, we can see that the numbers of errors of the type A (articulation errors) is not different in the three compared groups. However, comparing the increase of absolute numbers of speech errors of the type B, which refer to the prosodic features of speech, we found a greater increase ($P < .01$) in the schizophrenic group than in the both other compared groups.

We have found no statistically significant differences among the compared groups of examined persons in the ratio of the numbers of type A and type B errors.¹

¹ Analysis of variance and Duncan's multiple range test were used for the statistical checking of our data.

Because the difference between the schizophrenic patients and controls in the speech errors concerning the prosodic features was the most interesting finding, we asked whether this increase is not due to the difference which is present already without DAF. The mean number of speech errors of this type is 1,08 in controls, 1,21 in schizophrenic patients and 2,00 in LSD induced state. The differences are not statistically significant. Also the correlations between the numbers of errors without DAF and under the DAF conditions are very low and not statistically significant. We can conclude therefore, that the DAF method is necessary to provoke the greater increase of the type B speech errors in schizophrenics than in controls.

It is not easy to interpret the meaning of our results. We can speculate only on the character of the prosodic features of the speech where we found the difference: it seems that these features are in closer connection with emotional life than the region of articulation. It would mean that the emotional response of a schizophrenic patient to the intrusion induced by DAF differs from the response of a mentally healthy subject. The analysis of the type of the speech errors is then a useful method for elucidating such differences.

Numbers of examined persons

	Controls	LSD	Sch	Σ
men	25	18	54	97
women	40	19	81	140
Σ	65	37	135	237

Average numbers of speech errors

	N	A	B	A + B	A/B
Controls	65	4	22.36	27.41	0.11
Sch	135	4.31	39.61	38.97	0.20
LSD	37	3.87	25.90	27.86	0.08

REFERENCES

- Atkinson, C. J.: Adaptation to delayed sidetone. *J. Speech Hear. Disord.* 18: 386—391, 1953.
 Baštecký, J.: Delayed auditory feedback in psychiatry. Cand. Dissertation, Charles University, Prague 1965.
 Baštecký, J., Vinař, O.: Psychotropic drugs on delayed auditory feedback in schizophrenia. *Activ. Nerv. Super.* (Praha), 8: 222, 1966.

- Black, J. W.: The effect of delayed sidetone upon vocal rate and intensity. *J. Speech Hear. Disord.* 16: 56—60, 1951.
- Fairbanks, G.: Systematic research in experimental phonetics. I. A theory of the speech mechanism as a servomechanism. *J. Speech Hear. Disord.* 19: 133—139, 1954.
- Foulds, G. A.: A quantification of diagnostic differentiae. *J. Ment. Sci.* 108: 389—405, 1962.
- Goldfarb, W., Braunstein, P.: Reaction to delayed auditory feedback in schizophrenic children. In: *Psychopathology of communication*. New York, Grune and Stratton 1958.
- Hughes, F. W., Forney, R. B., Gates, P. W.: Performance in Human Subjects under Delayed Auditory Feedback after Alcohol, a Tranquilizer (Benzquinamide) or Benzquinamide-Alcohol Combination. *J. Psychol.* 55: 25, 1963.
- Koževnikov, V. A., Chistovich, L. A.: Реč, Артикуляција и Воспријатие. Изд. Наука. Москва, 1965.
- Lee, B. S.: Some effects of sidetone delay. *J. Acoust. Soc. Amer.* 22: 639—640, 1950.
- Linsener, J., Linsener, H. J.: Untersuchungen zum Lee Effekt I. Zur Klassifizierung von Sprachfehlern bei verzögerter Rückmeldung der Lautsprache. *Z. Psychol.* 168, 1—2, 26—58, 1963. Untersuchungen zum Lee Effekt II. Sprachstörung und Syntax. *Z. Psychol.* 168, 3—4, 226—230, 1963.
- Morávek, M.: Vliv zpožděné akustické aferentace (Leeův efekt) ve výzkumu vyšší nervové činnosti. *Activ. Nerv. Super.* (Praha) 3: 21—24, 1961.
- Spear, F. G., Bird, R. L.: Delayed auditory feedback. Vocal intensity changes in schizophrenia. *Brit. J. Psychiat.* 109: 240—242, 1963.
- Sutton, S., Roehrig, W. C., Kramer, J.: Delayed auditory feedback of speech in schizophrenics and normals. *Ann. New York Acad. Sci.* 105, 15: 832—844, 1964.
- Vinař, O., Baštecký, J., Borovičková, B., Zichová, M., Malač, V.: Method of delayed auditory feedback in psychiatry. *Activ. Nerv. Super.* (Praha), 7: 193—195, 1965.
- Vinař, O., Baštecký, J., Borovičková, B., Zichová, M., Malač, V.: Delayed auditory feedback in schizophrenia and LSD induced state. *Activ. Nerv. Super.* (Praha), 7: 302—303, 1965.

DISCUSSION

Singh:

- (1) Why did you delay the feedback by 0.142 and not by 0.18—0.20 which have been reported maximally deleterious to speech production?
- (2) Then you disagree with the previously reported data by Lee, Black, Fairbanks etc. that the maximal distortion takes place with 0.18—0.20 delay.

Vinař:

ad 1. We chose the delay of 0.142 sec because due to our previous experiences in a pilot study we could find maximal changes in parameters measured in our laboratory in the studied persons with this delay.

ad 2. We assume that disagreement with quoted authors is not a great one. For instance Black (1951) in one of his papers found that prolongation of time needed for reading under DAF condition increased as a function of the length of the delay until 0.18 seconds. These results were confirmed by Atkinson (1953) and Fairbanks (1955). The optimal interval of delay might be between 0.18—0.20 but in our experience the delay of 0.142 is sufficient which is not an exception (Sutton, Roehrig and Kramer 1964).

ÜBER DAS ERKENNEN SYNTHETISCH ERZEUGTER VOKALÄHNLICHER KLÄNGE DURCH NORMALSPRECHENDE, STAMMELNDE UND LESE-RECHTSCHREIBSCHWACHE SCHÜLER

RUTH BECKER*

Unter den Stammelern befindet sich erfahrungsgemäß eine Anzahl von Kindern, deren Sprachstörung vor allem durch die mangelhafte Unterscheidungsfähigkeit von Lauten begründet ist. Seeman (1) erklärt diesen Mangel damit, daß das kortikale Ende des Höranalysators ungenügende Differenzierungsfähigkeit besitzt. Selbst wenn diese Kinder einzelne Laute sprechmotorisch allmählich zu bewältigen lernen, differenzieren sie diese noch lange nicht mit dem Gehör. Der Autor äußert die Ansicht, daß eine sensorische Dyslalie stets mit einer Dysgraphie verbunden sei.

Schilling (2) verweist auf die verschiedensten Varianten des sensorischen Stammelns, die von der akustischen Unaufmerksamkeit, der mangelhaften psychischen Verarbeitung der Höreindrücke und einer Schwäche der auditiven Gedächtnisspanne bis zu echten Defekten der akustischen Diskriminationsfähigkeit aus dem Formenkreis der zentralen Hörstörungen reichen. Als besonders beachtenswert erscheint ihm die sogenannte partielle Lautagnosie — die er als fruste Form der akustischen Agnosie betrachtet —, die seiner Meinung nach oft nicht bemerkt wird. Diese Störung betrifft häufig nur einzelne Laute, die klangverwandt sind und nach einem ähnlichen Prinzip gebildet werden. Schilling erklärt diese Fehlleistung der Kinder als isolierte apperzeptive Störung, als Auffassungs- und Unterscheidungsschwäche für phonetische Klanggestalten.

Da sich die Artikulationsschwierigkeiten der Stammerler hauptsächlich im Bereich der Konsonanten äußern, konzentrieren sich die diagnostischen und therapeutischen Verfahren vorwiegend auf diese Lautgruppe. Dabei bleibt die Frage offen, ob Vokale von stammelnden und lese-rechtschreib-schwachen (Irs) Kindern ebenso gut wie von normalen Kindern gleichen Alters erkannt und differenziert werden oder ob es dabei Unterschiede gibt. Beim Erlernen des Schreibens (Rechtschreibens) werden bekanntlich sehr häufig Vokalauslassungen und -verwechslungen beobachtet.

Ziel der Untersuchung: Unsere vergleichende Untersuchung hatte deshalb das Ziel festzustellen, wie synthetisch erzeugte vokalähnliche Klänge von normalen, stammelnden und Irs Schülern erkannt werden. Wir wählten die am einfachsten

* Aus dem Institut für Sonderschulwesen der Pädagogischen Fakultät der Humboldt-Universität zu Berlin, Direktor: Prof. Dr. habil. K.-P. Becker.

erscheinende Form der Prüfung, indem wir die Vokale nicht im Kontext, sondern isoliert anboten. Dadurch ist es möglich, lediglich die Lautauffassung zu beurteilen, die als Voraussetzung für die analytische Bearbeitung eines Wortes unerlässlich ist.

Versuchspersonen (Vpn): Als Vpn standen je 30 normale und stammelnde Schüler aus 2. Klassen und 36 Irs Kinder aus zwei 2. und einer 3. LRS-Spezialklasse zur Verfügung, wovon 34 Kinder älter als die Kinder der beiden erstgenannten Vergleichsgruppen waren.

Wegen des Versagens beim Erlernen des Lesens und Rechtschreibens waren die Irs Kinder meist mehrmals sitzengelassen und hatten keine ihrem Alter entsprechende Klassenstufe erreicht.

Eine völlig altersgleiche Gruppierung der Kinder zu erhalten, scheiterte daran, daß

1. ältere Stammler auf Grund der systematischen Früherfassung und Frühbehandlung selten zu finden sind,
2. die LRS erst in 2. Klassen mit Sicherheit diagnostiziert werden kann, so daß sich bei den Irs Kindern diese Überalterungen ergeben.

Deshalb sind die drei Versuchskindergruppen nur bedingt vergleichbar. Alle Kinder waren audiometrisch überprüft worden.

Untersuchungsmaterial und Untersuchungsmethode: Uns stand ein Testtonband¹ von Janota (Prag) zur Verfügung, mittels dessen den Vpn 200 synthetisch erzeugte vokalähnliche Klänge angeboten wurden, die die Hörer als lange Vokale auffaßten. Die Klänge enthielten jeweils 4 Formanten, wovon der 1. Formant im Bereich von 294 Hz — 831 Hz und der 2. Formant im Bereich von 660—2349 Hz in Ganztonschritten verändert, der 3. und 4. Formant jedoch konstant bei 2500 bzw. 3500 Hz gehalten wurden. Die Kombination der Formanten F 1 und F 2, die sich bei jedem Vokal unterscheidet, wurde mit einer Impulsfrequenz von 110 Hz bzw. 220 Hz angeregt. Die Intensitäten der einzelnen Formanten wurden nach Fant eingestellt. (3) Der Abstand zwischen 2 Lauten betrug 5 Sekunden. Die Klangprodukte waren dadurch nicht nur als „Normallaute“, sondern in den verschiedenen Abstufungen als „Individuallaute“ zu hören, wie sie beim Sprechen im Wort wahrzunehmen sind. Die Kinder wurden aufgefordert, die gehörten, den Individuallauten vergleichbaren Klänge den Vokalen *a e i o u ä ö ü* zuzuordnen. (Vor die gleiche Leistung sind die Kinder beim Diktatschreiben gestellt.)

Die Zuordnung erfolgte in 3 Versuchen durch

1. Benennen der Buchstaben
2. Zeigen der Buchstaben
3. Schreiben der Buchstaben

Ein Vorversuch erfolgte an Erwachsenen und Schulanfängern.

¹ Herr Prof. Dr. habil. Lindner, Institut für Phonetik und Kommunikationsforschung der Philosophischen Fakultät der Humboldt-Universität zu Berlin stellte das Testtonband freundlicherweise zur Verfügung und beriet uns bei der Auswertung der Unterlagen.

Ergebnisse: Die Ergebnisse zeigten Übereinstimmung mit denen der Erwachsenen insofern, als auch bei den kindlichen Vpn Hörerurteile in Form von Konzentrationsfeldern für die einzelnen Vokale sichtbar wurden, die der Anordnung des Vokaldreiecks entsprechen (vergleiche Lindner). Dabei waren die Felder nicht immer scharf voneinander abgegrenzt, sondern zeigten Überlappungen. Die Felder wiesen allerdings bei den einzelnen Kindern eine sehr unterschiedliche Ausprägung auf und wichen auch zum Teil für die einzelnen Versuche stark voneinander ab. Erwartungsgemäß zeichneten sich die Unterschiede bei schwierig einzuordnenden Vokalen ab, die an der Grenze zweier oder mehrerer Vokalfelder stehen, z. B. *a — o*, *e — ä*, *ä — ör*. Vergleicht man die Hörerurteile bezüglich der Anzahl der Vertauschungen und ihrer Wertigkeit, so hat sich folgende Reihenfolge ergeben:

	Bestes Resultat	Mittleres Resultat	Schlechtestes Resultat
Sprechversuch:	norm. K.	Irs K.	stamm. K.
Schreibversuch:	norm. K.	Irs K.	stamm. K.
Zeigerversuch:	norm. K.	Irs K.	stamm. K.

Innerhalb der Kindergruppen zeigte sich eine unterschiedliche Reihenfolge:

	Normale K.	Irs K.	Stammelnde K.
Bestes Resultat:	Schreibversuch	Sprechversuch	Schreibversuch
Schlechtestes Resultat:	Sprechversuch	Schreibversuch	Sprech-/Zeigerversuch

Die meisten Vertauschungen erfolgten bei allen Kindern von *u* nach *o*, von *a* nach *o* und von *ä* nach *ö*. Die Stammler zeigten erwartungsgemäß die schlechtesten Resultate. Die Streuung der Urteile war bei den Irs Kindern in allen 3 Versuchen am größten.

Daher liegt es nahe anzunehmen, daß es sich bei den Irs Schülern um eine sehr heterogene Gruppe von Kindern gehandelt hat. Bei 60% der Irs Schüler war die isolierte Lautauffassung noch nicht ebenso sicher entwickelt wie beim Durchschnitt der Gruppe der normalen Kinder, obwohl sie fast alle älter waren als die letztgenannten.

Schlußfolgerungen: Die Ergebnisse belegen, daß die phonematische Differenzierung der Vokale stammelnden und Irs Kindern mehr Schwierigkeiten als normalen Kindern bereitet. Diese wirken sich beim Erlernen des Lesens und Rechtschreibens nachteilig aus.

Deshalb empfiehlt es sich, die Entwicklung und Schulung der phonematischen Differenzierungsfähigkeit nicht nur an Konsonanten zu betreiben, sondern auch die Vokale in die Übung einzubeziehen.

LITERATUR

1. Seeman, M.: Sprachstörungen bei Kindern. VEB Verlag Volk und Gesundheit Berlin, Jena, 1965, S. 103 ff.
2. Schilling, A.: Neuere Gesichtspunkte in der hals-nasen-ohren-ärztlichen Behandlung von Sprachstörungen. *Z. f. Laryngologie, Rhinologie, Otologie und ihre Grenzgebiete* 43. Jahrg., H. 6, (1964).
3. Lindner, G.: a) Über die Auffassung experimentell veränderter Sprache. „*Die Sonderschule*“, 1960, H. 3.
b) Beurteilung synthetisch erzeugter vokalartiger Klänge durch deutschsprachige Hörer. *Z. f. Sprachwissenschaft, Phonetik und Kommunikationsforschung* 1966.

DIE PAUSE IM KONTEXT IN FREIER ERZÄHLUNG

SVEINN BERGSVEINSSON

Wir werden uns damit abfinden müssen, daß die Normierung der Pausen im Kontext ein Problem ist, das noch seiner befriedigenden Lösung harrt. Es liegt in der Natur der Sache, daß die Ganzheitsgestaltung einer syntaktischen Struktur mehr Realisationsmöglichkeiten bietet als eine Elementarstruktur, etwa eines Wortes. Ursprünglich sind die Pausen physiologisch bedingt und insofern auch heute noch, da es begrenzt ist, wie lange wir eine Wortfolge aussprechen können ohne Atem zu holen. Andererseits sind die Pausen in einem Kontext lange nicht alle Atempausen. Ich habe (Grundfragen 1941) Pausen von 3/100 Sek. Dauer gemessen. Es ist ganz klar, daß so kurze Pausen bis etwa 25/100 (Frau Dr. Gericke, Tonband 1959) normalerweise keine Atempausen sind.

Jedoch ist allgemein bekannt, daß man nach jedem lexikalisch betonten Wort in einem Satz (etwa bei einem langsamen Diktat) je eine Pause einlegen kann. Das ist aber ein Sonderfall sowie auch die stockende Rede oder eine künstlerische Diktion. Wie es von mir (1941) und von Frau *Gericke* (Beiheft 7, 1959) festgestellt worden ist, beruht der Sprachrhythmus auf mehreren Faktoren. Sie sagt wörtlich S. 20: „Die Pause ist eben nur ein Mittel der rhythmischen Gliederung neben anderen. Allerdings scheint von den drei Versuchspersonen und erst recht von den Nachsprechern dieses Mittel bei den meisten Sprechereinheiten (SE) angewandt worden zu sein.“ Bei neuer Durchsicht meines Materials bin ich auch zum gleichen Ergebnis gekommen. Als Beispiel sei Sprechtext II (ST II) erwähnt (Unterhaltungssprache). Von 64 SE sind nur 8 ohne meßbare Pausengrenzen festzustellen oder 12,5%. Wenn auch die Grenzen der SE seinerzeit auf auditivem Wege, d. h. nach einem subjektiven Verfahren festgelegt worden sind, so stellen sich auch diese 12,5% als potentielle Pausen heraus. Beim langsameren Sprechen etwa könnten an diesen Grenzen Pausen eingelegt werden. Somit kommen wir nicht umhin, die SE als überordnete rhythmische Glieder in unsere Erwägungen zu ziehen, während die untergeordneten Glieder, die Silben, dabei außer Betracht bleiben können.

Bevor wir uns aber der Beziehung des Rhythmus zur Struktur zuwenden, möchte ich hier noch eine andere Frage anschneiden: *das Verhältnis zwischen Satzakzent und Pausierung*. Es wäre von vornherein zu erwarten, daß die rhythmischen Einheiten höherer Ordnung, also SE, einen und nur einen Satzakzent tragen würden, (vgl. die Akzentgruppen, stress-groups von *Sweet*, *Sievers* und *Viëtor*). In unserem wieder-

holbaren Material finden wir aber SE mit mehr als je einem Satzakkzent. Dazu zwei Beispiele aus Beiheft 9 von *Gericke*. Der Sprecher war ein 30jähriger Lehrer in Leipzig:

— Wir wollen ja 'unser 'neues Opernhaus aufbauen.

— Es hat zwar 'einige 'Schwierigkeiten gegeben.

Warum hier je zwei Satzakkzente? Ein wichtiger Konstituent der Phrase als SE ist doch ein Hauptakkzent. Im ersten Satz würden wir am ehesten „Opernhaus“ betont erwarten, nicht „unser + neues“. Im zweiten Satz nur das Wort „Schwierigkeiten“, nicht aber das Attribut „einige“. Dies Abhörergebnis liegt m. E. teils in der Betonung des Redeabschnitts als Stilprinzip: Gegensatzbetonung, Nachdruck, Emphase. Auch kann der Sprecher — bewußt oder unbewußt, oder auch aus Befangenheit wegen der Aufnahme — Silben mit normaler, logischer Betonung auf gleiche Ebene stellen, will sagen *betonen*, wie die mit Nachdruck- oder Gegensatzwert. Daher müßte der Abhörer sein Material nach diesen Gefühlswerten schichten. Teils liegt es in dem subjektiven Abhörverfahren selbst mit der ebenso subjektiven Akzentschätzung. Ähnliches gilt übrigens auch für das strukturelle Reduktionsprinzip beim Wortakkzent (Bierwisch 1966).

Bei der Abstufung der Rede in SE muß die Akzentschätzung ganzheitlich vorgenommen werden, nicht nur durch Vergleich einzelner Silben. Es muß also beim Abhören eine Methode zugrunde gelegt werden, die auf einem Wahrnehmungsprinzip beruht: *dem Gewicht der zeitnächsten Perception*. Auf dieser Grundlage muß man auch die manchmal schwierige Wahl des Akzentzeichens treffen. Dieses Prinzip besagt, daß in einer SE mit zwei oder mehr annähernd gleich schweren Akzenten der letzte überwiegt, d. h. in einer Reihe immer der letztere den ersteren übertönt — wahrnehmungsgemäß betrachtet. Wenn wir dieses Prinzip beim Abhören berücksichtigen, so wie auch die schon erwähnte Stilwertung der Akzentuierung des Sprechers, erreichen wir weitgehend eine Harmonie zwischen Satzakkzent und SE und damit auch zwischen dem Hauptakkzent (HA) und der Pausierung. Die Rolle der Intonation ist hier der Kürze halber beim Akzent mit einbegriffen.

Wir wollen uns nun der wichtigen Frage des *Rhythmus und der Satzstruktur* zuwenden. Sind die rhythmischen Einheiten nach strukturellen Regeln gegliedert, oder ist die Pausierung zufällig? Stockende Rede ist außerlinguistisch zu erklären. Von strukturalistischer Seite fällt die Frage zusammen mit der der Phrase und ihren Grenzmerkmalen. Aber es fehlt noch die Definition der Phrase.

Es scheint sich hier um grundverschiedene Sprachmittel zu handeln: einerseits der Sprachrhythmus, der sich außer den Pausen verschiedener prosodischer Mittel bedient, andererseits die Hierarchie der Satzstruktur, die syntaktischen Regeln folgt. Und aus Erfahrung wissen wir, daß einer rhythmischen SE engere Grenzen gesetzt sind als einem Satze. In der Syntax hören wir von Satzgliedern, nicht aber von rhythmischen Einheiten. Wir wissen aber durch den Kommunikationseffekt, daß diese Diskrepanz nur scheinbar sein muß. Die Pausen können nicht zufällig in die Rede eingestreut werden, auch werden sie nicht nach den Regeln der Grammatik gesetzt. Begriffe wie Sprechereinheit und Phrase sind nützlich

müssen aber durch Experiment und Statistik in Einklang miteinander gebracht werden. Diesen Ausführungen liegen Untersuchungen an zwei isländischen und drei deutschen Texten zugrunde (Grundfragen 1941, Beihefte 7, 8, 9 von Gericke, Leipziger Sprechner). In überwiegender Mehrzahl bildet Hauptsatz oder Nebensatz eine SE, oft auch Satzteile ohne Subjekt. Im Nachsatz kann auch ein Verb ausfallen: — Ich habe im Faulenzer geschlafen ... und er auf der Ottomane (H. 7). Adverbiale als SE wie — im Jahr (H. 8) kommen öfter vor, auch mit Infinitum: — in Paunsdorf beschäftigt (H. 8). Beachtenswert sind die Adverbiale im Vorfeld: — Mal sonntags (vorm Kinderfest) da sagt er auch. Sie werden von einem Stützwort *da* begleitet. Das gleiche finden wir bei einer Subjekt-Kette: ... mein Vater, voller Stolz. (*der* erzählt nun). (Beide Beispiele aus H.7). Wiederholungen sind oft SE: — dorthin (dorthin) dorthin (H.8). — jeden Tag /jeden Tag (H.8) — (knieend/knieend) (H.8). Mehrwörtige Appositionen: — Solche kleinen Dinger (H.7). Zahlenangaben: — neun, halb zehn (Prädikativ H.7), — zehn, elf Jahre (Apposition, H.7). Volladverb: — sumpfig (H.9). Dazu kommen als SE vor: Ausrufe, Anreden, Modalwörter wie *na*, *also*, *nun*, besonders satzeinleitend, und *ja* und *nein*. in normaler freier Erzählung. Nach jeder SE ist eine Pause: z. B. *neun* (P) halbzehn.

Was im untersuchten Material nicht vorkommt, sind Ein-Wort-Einheiten vor Pause außer satzeinleitenden Füllwörtern. SE wie „sumpfig“ (H.9) könnte nicht an erster Stelle stehen. Ein Adverbial vor dem finiten Verb darf nicht kürzer als zwei Wörter oder drei Silben sein. Dann gibt es auch keine Pause zwischen Subjekt und Verb (Interdependenz) oder nach einem Partikel mit Rektion, etwa — von seiner Ottomane (H.7). Nur eine Ausnahme in gleichem Heft: — Die waren sicher unter das (Pause) Sofa gerutscht. Ein Fehlsprechen?

Das Verhältnis Rhythmus: Struktur weist also Gesetzmäßigkeiten auf. Rhythmische Sprechereinheiten bis zur 0-Pausengrenze kommen überall zu ihrem Recht im Sinne der Sprachverständigung ohne die Sprachstruktur zu stören. Nur die engsten Verbindungen (Abhängigkeiten, Dependenz) erlauben keine Pausen. Darüberhinaus sind Pausen zwischen Ein-Wort-Satzteilen, je nach ihrer Stellung, begrenzt. Bezeichnend ist, daß *mal sonntags* eine SE im Vorfeld bildet, nicht aber ein einzelnes Adverb. Das gleiche gilt für die Zahlwörter. Durch erhöhtes Tempo fallen die möglichen Pausen weg. Obligatorisch ist dann nur der selbständige Satz bzw. die Periode.

LITERATURVERZEICHNIS

- Bergsveinsson, S.: Grundfragen der isländischen Satzphonetik (Munksgaard, Kopenhagen 1941).
Bergsveinsson, S.: Über die Bedeutung und die Auswertung der Satzmelodie. *Arch. f. vgl. Phonetik*, Band 5 (1941).
Bierwisch, M.: Regeln für die Intonation deutscher Sätze. *Studia grammatica VII* (Berlin 1966).
Garvin, P. L.: On Linguistic Method (Der Haag 1964).
Gericke, Ingeborg: Beobachtungen zur Melodie der Leipziger Umgangssprache (Dissertation 1956 in Manuskript).

Gericke, Ingeborg: Beihefte zum Magnettonband (Berlin 1959), Nr. 7, 8, 9.
Hjlemslev, L.: Omkring sprogteoriens grundlaeggelse (Kopenhagen 1943).
Kiparski, P.: Über den deutschen Akzent. *Studia grammatica* VI (Berlin 1966).
Mahnken, Irmgard: Formelemente des Sprechrhythmus. *Z. Phonet.* (1953).

DISCUSSION

Meinhold:

Nach eigenen Beobachtungen gibt es keine obligatorischen Pausen. Jede Pause kann durch terminale Verzögerungen oder bestimmte Tonhöhenmarkierungen am Syntagmaschluß ersetzt, ja sogar vorgetäuscht werden; dies gilt sogar für Pausen zwischen den Sätzen. Insofern dürfte die Pause auch in rhythmischer Hinsicht ein weitgehend nebensächliches Phänomen sein, trotz der Häufigkeit, mit der sie beim Sprechen auftritt.

Zacharias:

Die Pausensetzung bzw. Phrasierung hängt unter anderem vom Sprechtempo ab. Bei sehr langsamem Sprechtempo kann die Einheit eines Satzes sogar in einzelne Aussprüche zerfallen (Lindner), allerdings nur an den sogenannten fakultativen, erlaubten Pausen.

Schubiger

fragt, ob nicht bei den unerwarteten zusätzlichen Akzenten (,einige' Schwierigkeiten etc.) der Einfluß der Radiosprecher mit im Spiel ist, welche dazu neigen, durch zusätzliche Akzentuierung ihre Rede eindrücklicher zu gestalten.

De Vriendt:

1. Die fünf analysierten Sprechtexte sind freie Erzählungen. Wäre das Verhältnis zwischen Rhythmus und Satzstruktur das gleiche gewesen, wenn es sich um spontane Gespräche gehandelt hätte?
2. Sind die untersuchten Pausen wahrgenommene Pausen oder oszillographisch festgestellte Pausen?

Bergsveinsson:

ad Schubiger: Ja, mit diesem Faktor habe ich schon gerechnet. Im genannten Fall ist aber auch ein Nachdruck auf ,einige...' ebenso wahrscheinlich, ein auf den Inhalt bezogener Akzent (betontes Sprechen). Schließlich können mehrere Faktoren zusammenwirken. Darüber soll der Abhörer Angaben machen. Die enge Beziehung zwischen Rhythmus und Satzstruktur zeigt auch das Gesetz der zeitnächsten Perception, daß nämlich der letzte Starkakzent in einer Wortgruppe die als eine rhythmische Einheit (Sprecheinheit) empfunden wird, als Hauptakzent (Satzakzent) wahrgenommen wird.

ad de Vriendt: 1. Da liegt kein grundsätzlicher Unterschied vor. Im spontanen Gespräch kann auch der eine Gesprächspartner minutenlang ununterbrochen reden. Ganz kurze Repliken könnten höchstens das Resultat ändern. Auch unterschiedliche Individualitäten, etwa rede-gewandt: nicht rede-gewandt. Sonst haben die Abhörergebnisse die festen und lockeren Fügungen der SE die Beziehungen zwischen Rhythmus und Syntax bestätigt.

2. Es wurden wahrgenommene Sprechereinheiten zugrundegelegt. In überwiegender Mehrzahl fallen diese Grenzen mit den Pausengrenzen zusammen. Pausen unter 10/100 Sek. waren nur kymographisch bzw. oszillographisch festzustellen. 0-Realisierungen sind lediglich potentielle Pausen.

TWO CRITICAL PROBLEMS IN ACCENT RULES

(Shortened version of the paper given at the Congress)*

MANFRED BIERWISCH

Since 1956 Chomsky and Halle have developed a theory of accent which accounts for the complex pattern of relative prominence in compound words and phrases by fairly simple rules. These rules are strictly local transformations, recognizing only branching and categorization of the final derived constituent structure as well as the previously assigned stem-accent. They are operating cyclically from the innermost constituents upward to more and more embracing constituents. Two critical problems of this theory are discussed.

Firstly it is shown that the correct determination of the accent pattern of arbitrary sentences requires a specification of the domain in which the accent rules are to operate. This domain cannot be given by particular categories such as NP, VP, Sentence, etc. It must be connected somehow with the specification of breath groups or phrasing units. No solution is available for the moment.

Secondly it is shown that the accent pattern of whole sentences, but also of noun phrases does not only depend on labelled bracketing, but also on deeper syntactic, in particular on contextually determined information. From this it follows that the basic assumption that accent rules operate cyclically recognizing only the final derived constituent structure must be modified somehow.

* The complete text appeared in *Journal of Linguistics* Vol. 4. 1968. p. 173—178.

TWO PROBLEMS OF SOUND PERCEPTION AND PHONEMIC INTERPRETATION IN POLISH

HENRIK BIRNBAUM*

1. /a/ vs. /o/

As an illustration of the fact that children (as well as foreigners) often fail to grasp a given phoneme's specific vocal-tract configurations while being perfectly able to internalize an adequate table of phonemes in terms of their sensory actualizations. Jakobson adduced in his 1966 Moscow paper on phonic elements and speech perception the case of a three-year-old Polish boy who would substitute /a/ for Polish /o/ in spite of his being able to perceive the difference accurately so that he would attempt to correct the mispronunciation [dapaćągu] as repeated after him (instead of [dopocóągu]) by insisting on what might seem a tautological contradiction: „One can't say [dapaćągu], one must say [dapaćągu]!”² This inability of properly distinguishing between the articulation of the phonemes /a/ and /o/ in Polish, while clearly perceiving the acoustic difference, can probably be related to the fact, recently emphasized by Stieber, that the acoustic effect of Polish /o/ as distinct from /a/ can be achieved in one of two ways: either by a slight labialization concomitant with a raising (and simultaneous retraction) of the back part of the tongue or by a more strongly marked performance of either one (but not both) of these two articulatory movements.² In terms of distinctive features as now reformulated for Polish by Jassem the difference in articulation thus achieved in one of two possible ways amounts to a difference in the low vs. high-tone feature: /o/ (along with /u/ and /w/) is a low-tone vowel while /a/ (as well as /i/, cf. below) is a high-tone vowel, this feature not being relevant for all Polish vowels (viz., not for /i/, /e/, /j/)³

2. /i/ and /i̯/

Among the many unsettled problems of synchronic Polish phonology is the phonemic interpretation of the relationship of /i̯/ and /i/ (the latter written *y*). According

* (UCLA).

¹ R. Jakobson, "The Role of Phonic Elements in Speech Perception" (forthcoming). In the conventional IPA transcription the palatal voiceless affricate is symbolized by [tʃ] instead of [tʃ̟].

² Z. Stieber, *Historyczna i współczesna fonologia języka polskiego*, Warsaw, 1966, p. 99.

³ W. Jassem, *Biuletyn Polskiego Towarzystwa Językoznawczego*, XXIV (1966), pp. 98 and 102

to Stieber these two sounds must be considered allophones of one and the same phoneme, at least for the time being and in the language of a vast majority of educated Poles.⁴ Jassem, on the other hand, classifies /i/ and /i̇/ as separate phonemes.⁵ As is well known, a similar discrepancy in phonemic interpretation continues to exist for the corresponding vowels of modern Russian.⁶ In terms of Jassem's distinctive features, Polish /i/ and /i̇/ differ as follows: /i/ is acute, long (as opposed to short /j/), and does not participate in the low/high-tone opposition; /i̇/ is grave, high-tone, and does not participate in the short/long opposition. There is no reason to question the accuracy of the measurements and instrumental data reported by Jassem. How, then, can these conflicting views (as represented by the two Polish scholars) possibly be reconciled? The key to accounting for the two seemingly opposite views lies in a correct appraisal of the underlying respective concept of the phoneme. Jassem (oriented primarily toward physiological data) seems to conceive phonemes as made up exclusively of their constituent distinctive features. Phonemes could therefore be considered merely space-saving, conventional symbols representing actual sets of specific feature complexes; cf. the bracketed feature specifications used in generative phonology. While not explicitly stated, Jassem's concept appears thus to be closely related to the view recently advocated by Chomsky and Halle denying the need, and indeed the possibility, of setting up an independent phonemic level of representation. Stieber, on the other hand, shares the view that a phoneme as a whole is characterized by more than merely its constituent features. Most important of these additional characteristics are a phoneme's distributional properties which can modify (neutralize) features otherwise considered distinctive. It is on the basis of the complementary distribution of Polish /i/ and /i̇/ (granting certain exceptions as possibly foreshadowing an imminent breaking away of /i̇/ as an independent phoneme) that Stieber maintains their allophonic relationship. While both Stieber and Jassem thus recognize a substantial difference in actual articulation and perception of the two sounds, this difference is not considered phonemically relevant by Stieber whereas for Jassem, who does not take into account distributional criteria but only views phonemes in auditory-articulatory and acoustic terms, it is sufficient to separate /i/ and /i̇/ as independent phonemes. Both views can be justified if we realize that we deal here with different levels, the introduction of distributional criteria implying a higher level of abstraction. It is largely also in these terms that one must see the current disputes concerning the very existence of phonemes as truly linguistic entities (instead of merely equalling the sum of their constituent features). Thus on the feature level (i.e., without recourse to distribution), Polish /i/ and /i̇/ must be considered as constituting different feature

⁴ Z. Stieber, op. cit., pp. 101—102.

⁵ W. Jassem, op. cit., pp. 87, 95, 97—98, 101—104, 106.

⁶ Cf., e.g., the arguments for the minority view that Russian [i] and [i̇] constitute separate phonemes in A. N. Gvozdev, *Izbrannyye raboty po orfografi i fonetike*, Moscow, 1963, pp. 98—101, and *passim*.

complexes. On the phoneme level, on the other hand, the feature oppositions ascertainable on the lower level are suspended by the criterion of complementary distribution rendering the features differentiating /i/ and /i̇/ perceptually and genetically predictable and hence non-distinctive.⁷

DISCUSSION

Mańczak:

I und y sind zwei Phoneme, weil alle Polen sie ohne jede Schwierigkeit isoliert aussprechen können, was im Falle der kombinatorischen Varianten nicht vorkommt. Ein phonetisch ungeschulter Pole ist nicht imstande, die Variante von a in *niania* oder die von *ń* in *pański* isoliert auszusprechen. Obwohl in den einheimischen Wörtern *i* im Anlaut nicht vorkommt, ist nicht wahr, daß es dort unaussprechbar ist, vgl. *ypsilon* oder scherzhaftes *ymynyny* „imieniny“. Keine Beweiskraft hat *igrek*, da frz. *i grec*.

Lüdtke:

Zur polnischen Metasprache: Wie bezeichnet der normale Sprecher die Laute [i] und [i̇] sowie die ihnen entsprechenden Grapheme *i* bzw. *y*?

Hamm:

On the distinctive feature level, which phoneme is or will (theoretically) be marked, *i* or *i̇* (and which unmarked)?

Birnbaum:

As for the question of markedness, raised by Prof. Hamm, I would think that, in traditional terms, [i̇] would have to be considered the marked member of the [i] ~ [i̇] phoneme, [i] being the basic variant of that phoneme. However, this question becomes irrelevant in terms of the distinctive features of these sounds, as now restated by Jassem who considers [i] and [i̇] different phonemes. Thus the distribution of features is, in his view, as follows:

[i]		[i̇]
acute	vs.	grave
long (marked)	vs.	∅ (unmarked)
∅ (unmarked)	vs.	high-tone (marked)

Was die Benennung des [i̇]-Lautes im Polnischen betrifft, so heißt er dort entweder „y“ (d. h. [i̇]) oder „igrek“ (aber nicht „ygrek“!). Indessen ist die Benennung dieses Lautes kaum entscheidend für die Frage seiner phonematischen Selbständigkeit. Wichtiger ist die für die Verteilung (Distribution) wesentliche Tatsache, daß im Polnischen kein Wort mit „y“ beginnt. Natürlich werden [i] und [i̇] von linguistisch nicht vorgebildeten Polen stets als verschiedene Laute

⁷ For a summary of the recent differences of opinion concerning the existence of phonemes as independent linguistic units and for a brief discussion of the insistence on different levels of abstraction (as advocated by Malmberg, Šaumjan, and others), see my paper „Syntagmatische und paradigmatische Phonologie“, *Phonologie der Gegenwart*, Graz—Vienna—Cologne, 1967, pp. 307—352.

empfunden, was jedoch wiederum nicht unbedingt als Beweis dafür angeführt werden kann, daß diese beiden Laute nicht auf einer höheren Abstraktionsebene, nämlich gerade der phonematischen (wenn man an der Annahme einer solchen Zwischenebene trotz der bekannten Einwände der generativen Phonologie festhalten zu dürfen glaubt), als Allophone (Kombivarianten) desselben Phonems gewertet werden könnten. Daß die Frage, ob poln. [i] und [i̯] als selbständige Phoneme oder als Allophone des gleichen Phonems gewertet werden sollen, weitgehend von der Deutung des Phonembegriffs abhängt, habe ich ja in meinem Referat zu zeigen versucht.

THE MAGNITUDE OF PITCH INFLECTION

JOHN W. BLACK*

Pitch inflection here is intended to include any noticeable change in the frequency of the glottal tone during phonation. The change may be upward or downward, rapid or slow, and of greater or lesser extent. Specifically, the purposes of the investigation reported here were (a) to compare the impact of pitch inflections of two contrasting directions, extents, and durations or rates, and (b) to compare four procedures for evaluating this impact.

Procedure. Twenty university students served as listeners throughout the experiment. They responded to two sets of stimuli. Set One was extracted from recordings that Cowan made of a lengthy speech by the actor, Kenneth MacKenna. The speech was one of his lines in a then current play in New York City. Cowan determined the average fundamental frequency of successive $1/24$ second, and plotted the results graphically with the musical scale as the ordinate and time as the abscissa. This melody curve is illustrated in Figure 1.

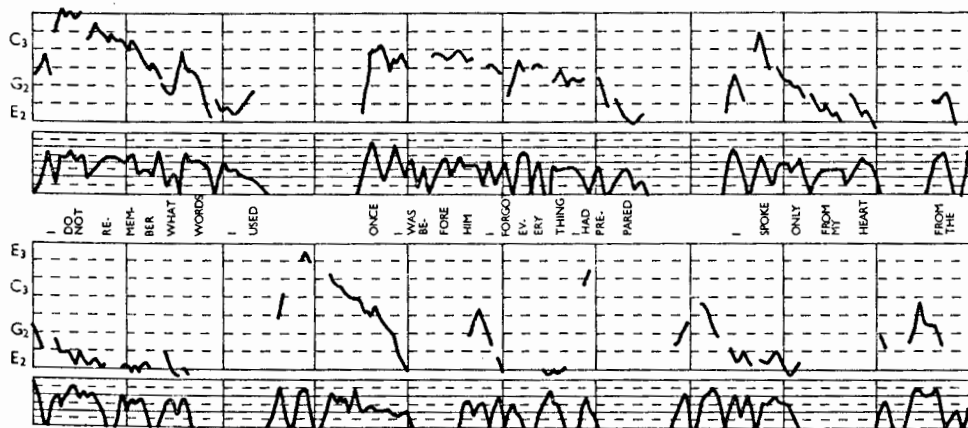


Figure 1. Melody curve from Cowan, "Pitch and Intensity Characteristics of Stage Speech", 1935.

* Ohio State University, Columbus, Ohio.

The materials of Set Two were borrowed from Takefuta who asked students to read sentences in manners appropriate to specified situational contexts, amazement, fear, indifference, etc. The vowel inflections were analyzed by a Trans Pitchmeter and photographed on an associated oscilloscope. The outcome is illustrated in Figure 2.

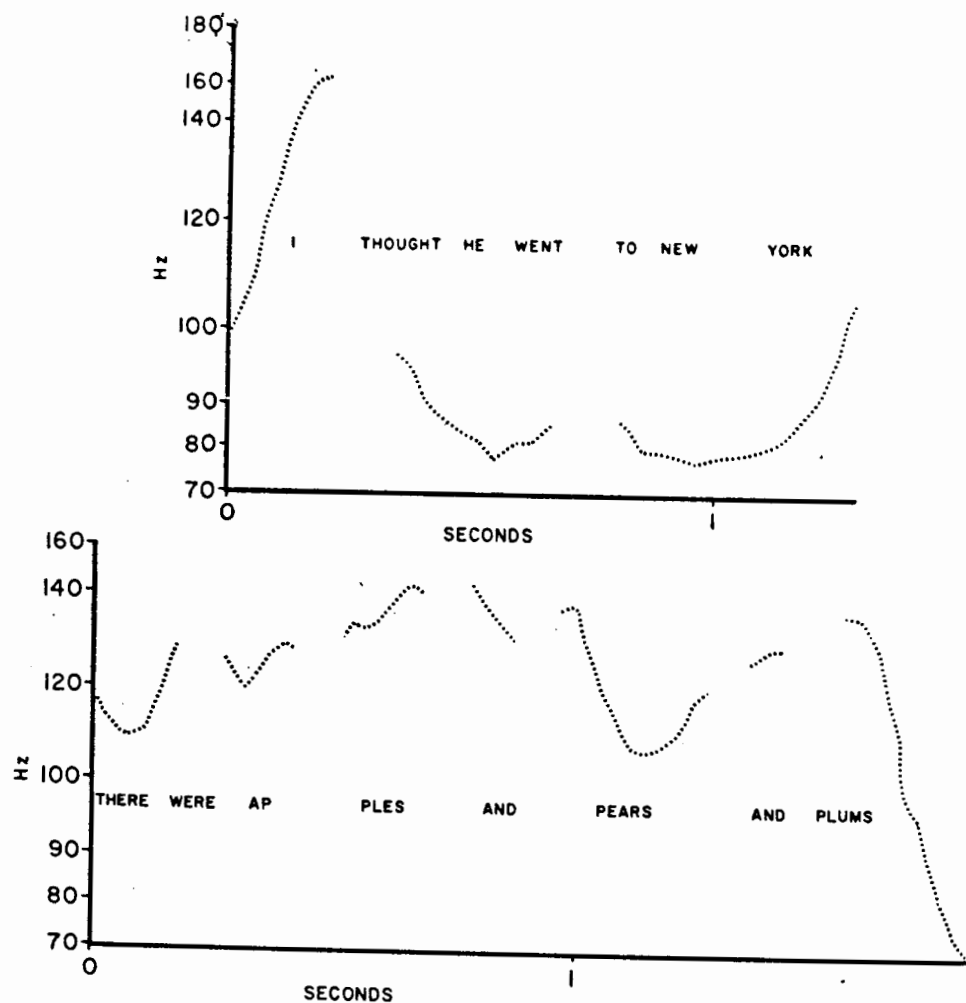


Figure 2. Display from oscilloscope of vocal inflections as analyzed by a direct-reading fundamental frequency instrument, the Trans Pitchmeter. Read-out on oscilloscope from a Trans Pitchmeter.

Insofar as possible, the 32 stimuli selected from Set One were matched by 32 stimuli selected from Set Two. These are summarized in Figure 3. There were four stimuli of each category listed in Figure 3. Thus there were four in the upward direction, relatively small in extent (3—5.5 semitones), relatively fast, i.e. short in duration (.10—.15 sec.). There were also four in the upward direction, of relatively great

extent (8—11.5 semitones), relatively slow, i.e. long in duration (.25—.40 sec.). The several permutations of these dimensions, upward and downward, are evident in Figure 3.

Category	(Cowan)		(Takefuta)	
	Extent (semitones)	Duration (sec.)	Extent (semitones)	Duration (sec.)
Up:				
Short-fast (✓)	5.0	.14	4.0	.10
Short-slow (✓)	4.0	.26	4.0	.30
Long-fast (✓)	9.0	.12	9.8	.20
Long-slow (✓)	8.0	.29*	10.0	.30
Down:				
Short-fast (✓)	4.0	.15	5.0	.10
Short-slow (✓)	4.0	.30	5.0	.40
Long-fast (✓)	9.0	.12	9.0	.15
Long-slow (✓)	9.0	.30	10.0	.30

* Median.

Fig. 3. Extent and duration of pitch inflections (modal values).

Listeners evaluated the stimuli on four separate occasions, using a different procedure each time: magnitude estimation, cross-modality matching, equal-appearing intervals, and paired comparisons. The salient features of the four procedures are evident in Figure 4. For *magnitude estimation* the subjects were asked to assign a num-

1.	<i>Magnitude Estimation:</i>	Assign a number; work in ratios.
2.	<i>Cross Modality Matching:</i>	Draw a line; work in ratios.
3.	<i>Equal-appearing Intervals:</i>	Encircle a value from 1—9.
4.	<i>Paired Comparisons:</i>	Answer—which is greater 1st or 2nd.

Fig. 4. Procedures.

ber to indicate the extent of the inflection of one segment and a number twice as large to indicate an inflection twice as great in extent. *Cross-modality matching* required the subjects to draw a line proportional to the extent of the inflection they heard. Again they were working in ratios. *Equal-appearing intervals* required the listeners to encircle a value from 1 to 9 on a previously prepared form. *Paired comparisons* required the listeners to answer the question: which is the greater inflection the first or second member of the pair?

Results. Numerical summaries of judgments of magnitude estimation, cross-modality matching, and equal-appearing intervals are directly comparable. These differ from the typical and most appropriate representation of outcomes of a paired-comparison procedure. Table 1 shows the median values of the judgments in the three directly comparable procedures. The values of magnitude estimation are in *ratios*; in cross-modality matching, *millimeters*; and in equal-appearing intervals, in units of a *scale from 1 to 9*. Irrespective of the psychophysical method, slow inflections were of greater impact than fast ones of the same extent, and upward inflections were generally more powerful than downward ones of the same extent and rate. (These generalizations are especially applicable to the Cowan material.) Long inflections were generally judged to be of greater magnitude than short ones of similar rate. The effect of short-slow inflections was quite comparable to that of long-fast ones. These outcomes, along with the results of the paired-comparison responses, are summarized grossly in Figure 5. The inflections of different extents and durations are arranged

Direction of Stimuli	(Cowan)				(Takefuta)			
	Cross Modality Matching	Method of Judging			Cross Modality Matching	Method of Judging		
		Magnitude Estimation	Equal-appearing Intervals	Paired Comparisons		Magnitude Estimation	Equal-appearing Intervals	Paired Comparisons
Up:								
Minimum	SF	SF	SF	SF	LF	LF	LF	LF
↓	LF	SS	SS	SS	SF	SS	SF	SF
Maximum	SS	LF	LF	LF	SS	SF	SS	SS
↓	LS	LS	LS	LS	LS	LS	LS	LS
Down:								
Minimum	SF	SF	SF	SF	SF	SF	SF	SF
↓	LF	LF	LF	LF	SS	SS	LS	SS
Maximum	LS	LS	SS	SS	LS	LF	SS	LF
↓	SS	SS	LS	LS	LF	LS	LF	LS

Fig. 5. Rank order of the relative effect of four types of vocal inflection: short-fast (SF), long-fast (LF), short-slow (SS), and long-slow (LS).

top to bottom, in an order of increasing impact. Thus, with upward inflections, the long-slow ones carried the greatest impact in all instances. A second purpose of the study was to examine the relative effectiveness of the alternative psychophysical procedures. In the instances of magnitude estimation, cross-modality matching, and equal-appearing intervals, the reliability of the listening panel was tested by an

analysis of variance as developed by Hoyt. In the instances of paired comparison, the reliability was tested by the split-half procedure. The outcomes are summarized in Table 2. Five of the eight values of *r* exceed .90. Although the present data do not justify a ranking of the four procedures the relative effectiveness of magnitude estimation is apparent.

Conclusion. The relative magnitude of an inflection appears to be measurable. The impact is greater with a larger inflection than a smaller one and with a slower one than a fast one, within the dimensions of the stimuli that were studied. Further, and of greater interest, the impact of an upward inflection is greater than that of a downward inflection when the two are matched in extent and rate. The magnitudes of inflections of different dimensions are ordered similarly whether one or another of the commonly used psychophysical procedures is employed.

REFERENCES

- Cowan, J. Milton, "Pitch and Intensity Characteristics of Stage Speech," *Archives of Speech* [Supplement], December, 1936.
 Takefuta, Yukio, "A Study of Relative Efficiency of Acoustic Parameters in the Intonational Signal of American English," unpublished doctoral dissertation, The Ohio State University, 1966.

DISCUSSION

J. T. Hart:

Your result, that under the same conditions of length, slope and extent, a rise is more effective than a fall, seems to be well in line with our finding that in a well balanced set of rise and fall for Dutch the slope of the fall has to be about twice as steep as that for the rise, whereas in that case the duration of the fall has to be at least three quarters of that for the rise.

Sadanand Singh:

I would like to ask two questions to Prof. Black.

- 1) He said "the upward inflexion had greater impacts". I did not understand—on what? And what are the implications of such a distinct finding.
- 2) Whether Takefuta replicated the study by Cowan, i.e., did he control variable, including experimental materials, similarly as Cowan or whether he had introduced new variables and used different experimental materials.

Black:

The implications of the fact that an upward inflection is more potent per semitone per unit of time than a downward one are numerous. They relate to effective voices (Black, 1942), to effective reading of literature, to the teaching of second languages, and to sociolinguistics. The Takefuta and Cowan materials are unlike except that both have been *measured*. They permit replications of studies such as this one.

PHONETIC PATTERNS OF CONSONANT SUBSTITUTION

IRENE BLANCHARD*

Precision of utterance of a language depends upon the exactness of the articulation of consonants. Adult American English speech uses 23 single consonant phonemes, combined with twice as many vowels in myriad combinations. These are the word codes for transmitting meanings. A child with intelligence, hearing, a good verbal example, and a well organized body, learns these sound combinations and codes in an orderly sequence from his first year to maturity. The acoustic signal does not operate alone. Firm establishment of consonants in his early vocabulary is accomplished at least by seven years of age. In the course of this normal development, when certain consonants may not yet have become firmly established, other consonants may be substituted for them in a pattern we call Baby Talk. Thus, if a three-year-old infant had reason to say "My little sister thought you were washing," it might come out instead: *mai jItə fItə sət u wə wosin* (loose transcription). The five year old with no front teeth would say: *mai ? Itl θItə θət ju wə wəθɪŋ*. These phonetic substitutions are to be expected during the early years of speech development. They may be called normal substitutions appropriate to the ages and conditions indicated.

Not so the phonetic substitutions of persons whose normal development has been aborted. Early difficulty in perception of high pitched phonemes would result in this pattern, sometimes prolonged into adulthood: *mə Itl tItə tət ju və wəθɪŋ*. A dysfunctioning palate might cause these phonemic alterations: *mə Ihə hIhə kə hu wə haki*. This effect happens after prolonged hypertrophied tonsils have been removed: *mə !Ilə !Ilə θlə u wə wəθli*. Ataxia, this: *mə Ikə kIkə kə ku aki*. Phonemic patterns in mental retardation are difficult to predict. This is one sample: *ma Ikə fIdə ha du ahi*.

Almost all beginning talkers say *dis* for *this*. This substitution of *d* for *ð* and the omission of *s* shows in the pattern of baby talk. These little children are likely to say *wItə wəbIt* until they go to school and read about *little rabbits*. The substitution of *w* for *l* and for *r* is appropriate in the Baby Talk pattern.

A word about American Adult Usage is in order. The stretch of our geography and the speed of our existence have led us to short cuts, even in energy expended for phonetic articulation! Some people call the result "careless speech". It may be care-

* Pacific State Hospital, Pomona, California U.S.A.

less teaching. Many of us substitute *n* for *ɳ* in the present participle, as *goin*, *kamən*. Progressively more of us call wheat *wit* and where *wet*, and what *wɔt*. And, especially in the west, we call garages *gəndʒəz*. The latter two substitutions have become so wide-spread in two generations of American speakers as to represent real phonetic change in American English speech.

Phonetically, we can trace the usual substitutions. Baby Talk gradually recedes from three to seven, the lisping substitution appears and disappears, and the adult pattern emerges. This sequence of development assumes that the children have enough intelligence and hearing to enable them to learn to speak. Hearing loss for high pitches results in sustained substitution of *t* for almost any phoneme. Mental retardation imposes bizarre patterns of phonemic substitution and omission. The great number of phonemes left out of the pattern is significant.

The normal pattern of substitution is appropriate to retarded children with IQ measurement of 70 and above. Below this point, however, and under circumstances critical to intelligent behavior, the patterns differ. Not only does the number of phonetic substitutions increase as the IQ becomes lower, but the character of the substitutions shows marked deterioration. The gradual imposition of *h* (sometimes *x* or *ç*) where it doesn't belong is unique to the retarded population.

Conditions that cause mental retardation contribute to the character of phonetic utterance. The familiar hearing loss that accompanies certain impairment cuts out from the verbal pattern all faint and high pitched sound, and peculiarly inserts the plosive *t* where it does not belong. Lack of stimulation, reward, or acceptable example, or inherited imperception of phonemic utterance contribute to undeveloped patterns of speaking. The peculiar sound of the mongoloid shows deviance from normal Baby Talk pattern, with much dependence upon *h*. So, also, does the utterance of those whose retardation results from traumatic accident. Other etiologic conditions produce distinctive patterns, and these patterns persevere through life.

Some patterns need not persevere. The patterns of children whose oro-pharyngeal structures are impeded by unaccustomed tissue growths can be corrected. And there is hope for the spastic child who cannot make himself understood. Simply by organizing his eating habits to develop a strong suck, as infants have, the spastic pattern is eliminated and spontaneous Baby Talk is heard. We can predict with security that the physical act of producing phonemes in coded verbal language depends upon basic functions that can be collected in four fundamentals, the foundation blocks of the pyramid of communication: intelligence, hearing, well-organized structures, and an environment conducive to speech, none of which act alone to produce speech. With this equipment, meaning develops which demands expression, first in gross physical action, then in melodies and stresses of voice and isolation of vowel values, in precision of utterance for phonemic articulation of words, and finally in complete language.

ИДЕНТИФИКАЦИЯ РЕЧЕВОГО СИГНАЛА В ЗАВИСИМОСТИ ОТ ЕГО ЗВУКОВОЙ СТРУКТУРЫ

Л. И. БЛОХИНА, Р. К. ПОТАНОВА*

Результаты исследований, проведенных с целью установления зависимости между надежностью идентификации речевого сигнала и его языковой структурой свидетельствуют о наличии такой зависимости. К числу лингвистических факторов, способствующих увеличению разборчивости речевого сигнала относят: а) частоту встречаемости отдельных слов (Н. В. Savin) б) увеличение количества слогов в анализируемом речевом сигнале (Н. В. Savin), в) принадлежность речевого сигнала к разным языковым уровням: фонологическому, морфологическому или синтаксическому (Л. Р. Зиндер, Л. И. Поварков, Г. В. Глекин), г) место ударения в слове (И. Луцкихина).

В настоящем докладе обсуждаются результаты, полученные в ходе исследования характера и степени влияния лингвистических факторов на надежность различения двух сигналов, идентичных по звуковому составу, но произнесенных или двумя разными дикторами, или одним и тем же диктором.

В качестве экспериментального материала были использованы звуки, слова и фразы, подобранные с учетом следующих факторов:

1. частотность употребления отдельных звуков, слогов и слов;
2. количество слогов в слове и фразе (длительность сигнала);
3. место ударного слога в слове;
4. спектральная характеристика звуков речи, входящих в состав экспериментальных слов и фраз.

Дикторами были 32 студента (мужчины, русские, москвичи, в возрасте от 19 до 23 лет).

Каждый диктор произносил список экспериментальных звуков, слов и фраз четыре раза. В результате монтажа ферромагнитной пленки были получены 64 пары сигналов. Порядок следования пар сигналов, произнесенных одним и тем же диктором, и пар сигналов, произнесенных разными дикторами, случайный.

* Москва, ИМГиШИА им. М. Горького, ЛЭФИИР.

Для аудирования полученного экспериментального материала были привлечены:

- а) 5 опытных фонетистов (возраст от 28 до 33 лет)
- б) 28 студентов (не выступающих в качестве дикторов) (возраст от 18 до 22 лет).

Прослушивание экспериментального материала проводилось бинаурально через наушники и через динамик.

В результате аудирования было получено 580 первичных протоколов. Все протоколы были обработаны по правилам вариационной статистики.

Анализ полученных данных свидетельствует о правомерности гипотезы, выдвинутой авторами настоящего доклада. Результаты исследования обнаруживают достаточно четко выраженную корреляцию между процентом правильно идентифицированных пар сигналов и такими факторами как: 1. длительность сигнала, 2. принадлежность сигнала к различным языковым уровням, 3. спектральная характеристика речевого сигнала.

Первым фактором, способствующим увеличению надежности идентификации является увеличение длительности сигнала. На каждом из двух высших иерархических языковых уровней (морфологическом и синтаксическом) наблюдается положительная корреляция между процентом правильно идентифицированных пар сигналов и длительностью соответствующей языковой структуры (см. график № 1).

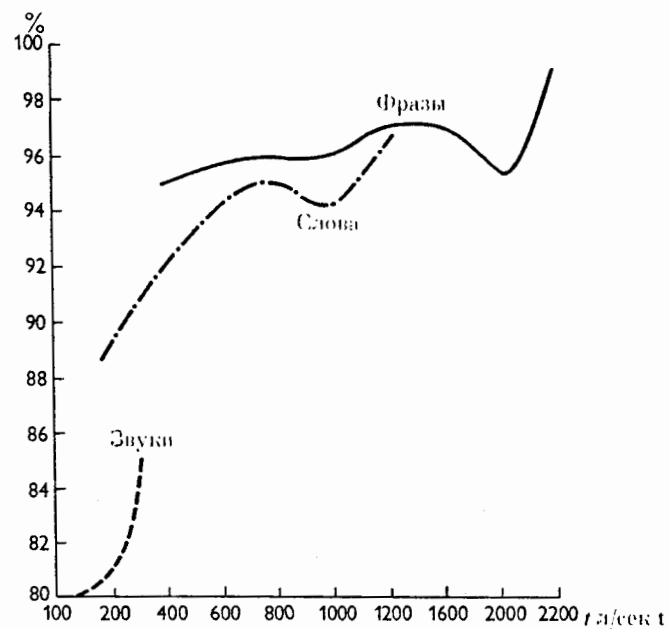


График 1. Зависимость процента правильных идентификаций от принадлежности сигнала к определенному языковому уровню и от длительности данного сигнала.

Что касается второго фактора — принадлежность сигнала к определенному языковому уровню — то, как показал подсчет правильно идентифицированных пар речевых сигналов, наиболее резкий скачок отмечается при переходе от фонологического уровня к морфологическому уровню (с 80 % правильно идентифицированных пар до 89 %).

Переход от морфологического уровня к синтаксическому выражен не так ярко, однако, и здесь во фразах отмечается увеличение процента правильно идентифицированных пар сигналов по сравнению со словами, имеющими то же количество слогов (см. график № 1).

Увеличение надежности идентификации речевых сигналов, принадлежащих к морфологическому и синтаксическому уровням, по сравнению с сигналами, принадлежащими к фонологическому уровню, следует объяснить, по нашему мнению, тем, что в распоряжении аудиторов имеется дополнительная информация о произносительных особенностях диктора, находящих свое выражение на морфологическом уровне в переходах от гласных к согласным и от согласных к гласным, в распределении произносительной энергии в слове, в особенностях выделения ударных слогов, а на синтаксическом уровне дополнительно в интонационном оформлении фразы.

Третьим фактором, влияющим на надежность идентификации речевых сигналов, является спектральная характеристика звуков речи, точнее

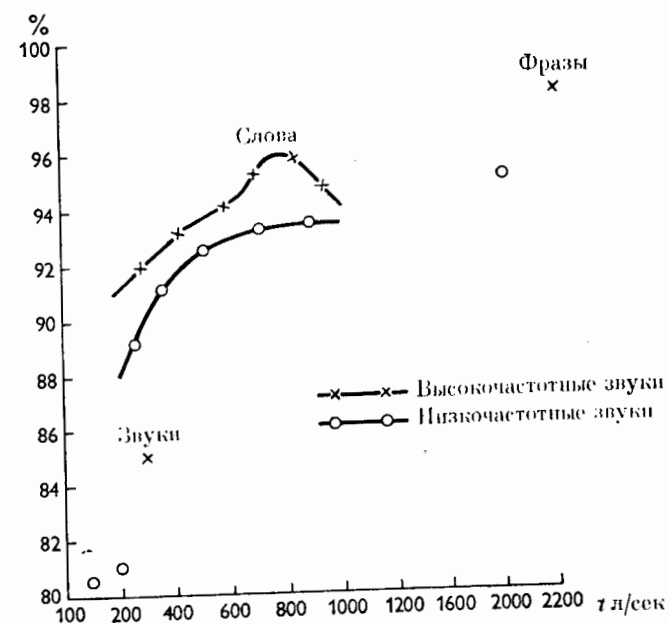


График 2. Зависимость процента правильных идентификаций от звукового состава сигнала и длительности.

место сосредоточения основной энергии в спектре звука (в области высоких или низких частот).

На всех языковых уровнях речевые сигналы, состоящие из „высоко-частотных“ звуков (фрикативные согласные, гласные переднего ряда) всегда давали более высокий процент правильной идентификации по сравнению с речевыми сигналами, состоящими из „низкочастотных“ звуков (сонорные согласные, гласные заднего ряда). (См. график № 2.)

Подобная зависимость должна быть объяснена, с нашей точки зрения, явлением, известным в акустике под названием маскировки. Маскировка, представляющая собой явление, обратное анализу, т. е. разложению акустического стимула на составляющие компоненты, затрудняет выделение слухом отдельных компонентов речевого стимула. Сильное маскирующее действие низких тонов подчеркивается рядом исследователей (С. В. Ржевский, С. С. Стивенс, Г. А. Миллер и др.). По данным вышеназванных авторов наибольшая маскировка вызывается тонами с частотой 300 гц—500 гц, т. е. теми частотами, в области которых сосредоточена основная энергия сонорных согласных и гласных заднего ряда.

Естественно предположить, что при сравнении двух речевых сигналов аудиторы вычленили из сложного акустического комплекса отдельные компоненты с целью их сопоставления. Явление же маскировки, имеющее место в „низкочастотных“ сигналах, затрудняло это выделение и тем самым снижало процент правильной идентификации речевых сигналов.

В отличие от опытов по разборчивости такие факторы как место ударения в слове и частота встречаемости отдельных языковых единиц не оказали определенно выраженного влияния на надежность идентификации речевых сигналов предъявленных аудиторам.

Сопоставление слов с разным местом ударения (‘ездили, объ’ездил, пере’езд) не дала однозначных результатов. В одной группе аудиторов на всех трех словах было получено одинаковое количество правильных ответов, в другой группе — слова с ударением на втором и третьем слоге дали несколько больший процент правильных ответов по сравнению со словами с ударением на первом слоге. Следует, однако, подчеркнуть, что процент правильных ответов увеличивается с введением в речевой сигнал второстепенного ударения (контрмеры).

Аргументом против значимости фактора частоты встречаемости слов для надежности идентификации сигналов в условиях нашего эксперимента являются данные, полученные при идентификации, с одной стороны, слов с большой частотой встречаемости (типа „электрификация“), с другой стороны, редко встречающихся слов (типа „диаэрация“). Для обеих групп слов был получен одинаковый процент правильных ответов.

THE APPLICATION OF INFORMATION THEORY TO VOWEL-RECOGNITION EXPERIMENTS

J. G. BLOM*

This paper deals with the application of Information theory to the transmission of natural and artificial vowels. I want to start with a brief explanation of the main concepts of information theory for those not acquainted with them.

Information theory describes the phenomena of transmission as perceived by an outside observer who has full knowledge of both sides of the transmission channel.

The symbols to be coded by the transmitting part of the channel—in our case in sounds—will be referred to as input, the symbols decoded at the receiving end as output.

In this case we are only interested in the most simple situation in which the auto-correlation of the string of input symbols is zero, which means that the input symbols are in a random order. Incidentally this does not exclude the possibility that the decoding process is affected by the actual succession of two or more sounds. The number of different symbols will be finite.

Due to imperfections or instability of transmitter and receiver, distortion or interference, the string of output symbols will not be an exact replica of the string of input symbols. We speak therefore of a transmission channel with noise.

Let the number of different symbols be n .

The symbols can then be referred to as $S_1, S_2 \dots S_n$.

The performance of the channel can be depicted by a table of confusion probabilities (see fig. 1).

$$\sum_j p_{j0} = \sum_k p_{0k} = \sum_j p_{jk} = 1$$

For a noise-free channel

$$\begin{array}{ll} p_{j0} = p_{jk} = p_{0k} & \text{for } j = k \\ p_{jk} = 0 & \text{for } j \neq k \end{array}$$

For a channel with no correlation between input and output (that means no transmission at all, the receiver is only guessing)

$$\begin{array}{ll} p_{jk} = p_{j0} \cdot p_{0k} & \text{for } j = 1, \dots, n \\ & k = 1, \dots, n \end{array}$$

* University of Amsterdam, Institute of Phonetic Sciences.

Input		Output			
		S_1	S_2	S_k	S_n
S_1	p_{10}	p_{11}	p_{12}	p_{1k}	p_{1n}
S_2	p_{20}	p_{21}	p_{22}	p_{2k}	p_{2n}
S_j	p_{j0}	p_{j1}	p_{j2}	p_{jk}	p_{jn}
S_n	p_{n0}	p_{n1}	p_{n2}	p_{nk}	p_{nn}
Total	1	p_{01}	p_{02}	p_{0k}	p_{0n}

Fig. 1. Confusion Probability Matrix.

p_{j0} = probability that S_j is the input symbol,
 p_{0k} = probability that S_k is the output symbol,
 p_{jk} = probability of the combination of S_j as input symbol and S_k as output symbol.
 In a forced-choice situation

A real channel will be somewhere between these extremes. Now we have to deal with different amounts of information. The information of the input H_x , that of the output H_y and that of the combination of input and output H_{xy} .

The unit of information is called a Bit.

One bit is the amount of information contained in the answer to a question to which there are two mutually exclusive answers with equal probability of occurrence. Take for example the information contained in the position of a coin. So the amount of information in Bits is the minimal number of questions of the type just mentioned necessary to obtain full knowledge. The amounts of information can be easily calculated using the formulae of fig. 2.

$$H_x = \sum_j -p_{j0} \log p_{j0}$$

$$H_y = \sum_k -p_{0k} \log p_{0k}$$

$$H_{xy} = \sum_{j,k} -p_{jk} \log p_{jk}$$

Fig. 2.

When information is transmitted by the channel we have the following inequality

$$H_x + H_y > H_{xy}$$

This means that given the output and our knowledge about the confusion matrix, we can make a good guess at the input.

The relations between H_x , H_y , and H_{xy} can be shown in simple Wenn-diagrams. (See fig. 3.)

$$T_{xy} = H_x + H_y - H_{xy}$$

The cross-section between H_x and H_y is called the transmission T_{xy} .
 The physical meaning of the transmission is that part of the information of the input which we know when the output is known, in other words, the transmission is the information transmitted by the channel.

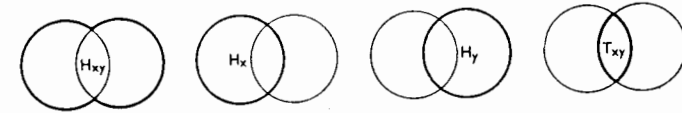


Fig. 3.

In order to calculate the transmission we have to make use of the confusion frequency matrix resulting from an experiment (fig. 4).

Input	Total	Output			
		S_1	S_2	S_k	S_n
S_1	m_{10}	m_{11}	m_{12}	m_{1k}	m_{1n}
S_2	m_{20}	m_{21}	m_{22}	m_{2k}	m_{2n}
S_j	m_{j0}	m_{j1}	m_{j2}	m_{jk}	m_{jn}
S_n	m_{n0}	m_{n1}	m_{n2}	m_{nk}	m_{nn}
Total	M	m_{01}	m_{02}	m_{0k}	m_{0n}

Fig. 4. Confusion Frequency Matrix.

In this table the m 's represent observed frequencies, the subscripts have the same meaning as in the probability matrix.

Taking the quotients m/M as best estimates for p 's we can calculate the transmission.

The necessary calculations can easily be programmed for evaluation by an electronic computer.

All our calculations were carried out with the IBM 1130 system of the Institute of Phonetic Sciences of the University of Amsterdam.

To get some insight into the process of vowel perception we applied information theory to some data published in the literature.

We started with the well-known experiment by Peterson and Barney on formant measurements on vowels of different speakers (JASA 1952) (fig. 5).

Suppose we have a vowel-recognition system that relates the sounds within a specific contour to one and only one vowel-class.

We determined the confusion frequency matrix for such a system shown in fig. 6 by a simple counting procedure, any sound falling in the cross-section of two areas being scored as 0.5 for each area. All frequencies are multiplied by 10 to avoid fractions.

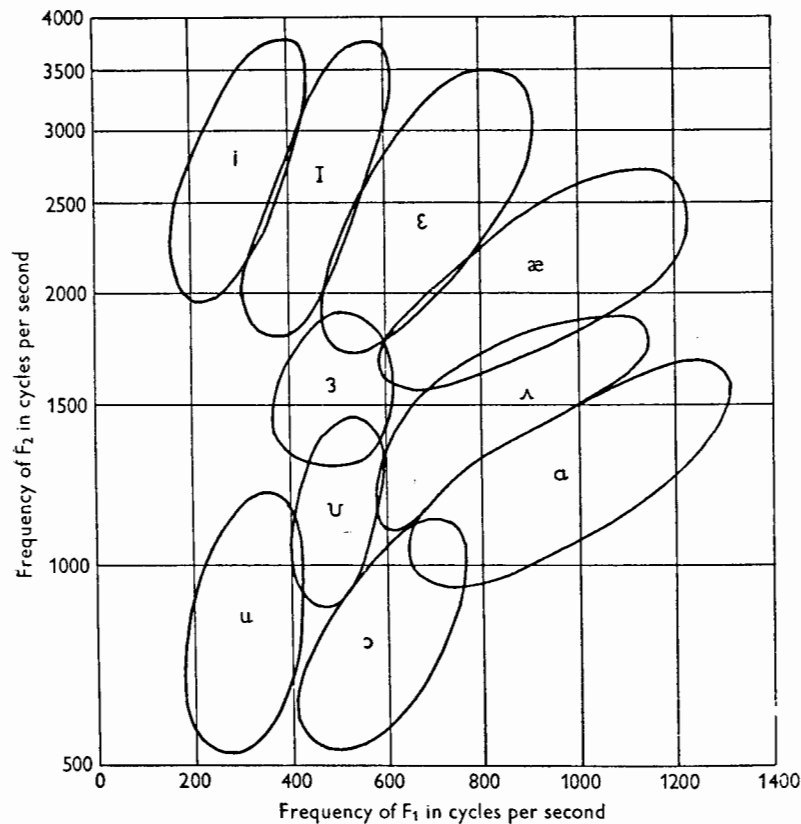


Fig. 5.

As we see, the information of the input is 3.32 Bits, the transmission 2.19 Bits. The same sounds were presented to a group of listeners. Peterson and Barney published the confusion matrix which is shown here as fig. 7.

When we apply our formulae to their matrix we find a transmission of 2.98 Bits.

It appears that human listeners do better than our hypothetical vowel recognition system. Our conclusion must be that man uses factors additional to the first two formants. These factors might be fundamental frequency, duration, the connection with surrounding consonants, and knowledge of the particular vowel system of an individual speaker. Although the speechsounds of different speakers were randomized, some knowledge of the position of the vowel system in the two-formant plane was available, due to the high correlation between fundamental freq. and the formant frequencies (Mol 1964).¹

As no confusion occurs when we listen to the sounds of a familiar voice we can list and add up our data as follows.

¹) Proceedings of the 5th Intern. Congress of Phonetic Sciences.

Peterson & Barney Formant Measurements											
Input	Total	Out-put 1	2	3	4	5	6	7	8	9	10
1	740	680	60	0	0	0	0	0	0	0	0
2	750	30	650	55	0	0	0	0	0	0	15
3	755	1	96	571	11	1	1	1	1	1	71
4	740	1	1	121	566	26	1	1	1	1	21
5	730	0	0	0	30	610	45	35	0	5	5
6	760	3	3	3	3	53	603	83	3	3	3
7	740	2	2	2	2	22	92	572	12	32	2
8	740	2	2	2	2	12	2	2	567	92	57
9	750	1	1	1	1	41	1	11	76	491	126
10	775	2	22	97	52	2	2	2	7	127	462
Total	7480	722	837	852	667	767	747	707	667	752	762

$H(X) = 3.32$ $H(Y) = 3.31$ $H(XY) = 4.45$ $T(XY) = 2.18$

Fig. 6.

Information of input		3.32 Bits
Contributed by formant positions alone	maximal	2.19 Bits
Contributed by other factors than system	at least	0.79 Bits
	Sum of these factors	2.98 Bits
Contributed by specific knowledge of an individual speakers vowel system		0.34 Bits
		3.32 Bits

The next data to be examined are published by Cohen, Slis & 't Hart (Phonetica 1967) in an article entitled "On Tolerance and Intolerance in vowel perception". They utterly failed to grasp the meaning of a paper by Blom & Uys, entitled "Some Notes on the Existence of a 'Universal Concept' of Vowels" (Phonetica 1966), but they presented a highly interesting confusion matrix for a system of 12 synthetic vowels. They used 12 fixed two-formant positions and introduced duration as an extra parameter. The spacing of the vowels in the F_1 , F_2 plane is somewhat exaggerated. The matrix is shown in fig. 8.

The information of the input is 3.62 Bits in formant positions and 1.55 Bits in duration which is redundant.

// XEQ
Peterson & Barney Listening Experiment

Input	Total	Out-put	1	2	3	4	5	6	7	8	9	10
1	10 280	10 267	4	6	0	0	3	0	0	0	0	0
2	10 279	5	9 549	694	2	1	1	0	0	0	0	26
3	10 277	0	257	9 014	949	1	3	0	0	0	2	51
4	10 278	0	1	300	9 919	2	2	0	0	0	15	39
5	10 273	0	1	0	19	8 936	1 013	69	0	228	7	
6	10 279	0	0	1	2	590	9 534	71	5	62	14	
7	10 279	0	0	1	1	16	51	9 924	96	171	19	
8	10 279	0	0	1	0	2	0	78	10 196	0	2	
9	10 277	0	1	1	8	540	127	103	0	9 476	21	
10	10 279	0	0	23	6	2	3	0	0	2	10 243	
Total	102 780	10 273	9 813	10 041	10 906	10 090	10 737	10 245	10 297	9 956	10 422	

$H(X) = 3.32$ $H(Y) = 3.32$ $H(Y) = 3.66$ $T(XY) = 2.98$

Fig. 7.

The transmission is 2.93 Bits. As the experimental conditions are comparable with the situation in which a person is listening to the sounds of one individual speaker, part of the information is lost. (Of course, some of the factors operating in experiments where monosyllabic words are used are absent in experiments with isolated sounds).

This low transmission is in agreement with our findings. It seems that a transmission channel operates less stably with artificial vowel-like sounds than with natural vowels.

From the results of the scaling experiment described by my colleague Meinsma an estimate can be made as to the confusion occurring between different areas of the perceptive vowel-triangle. We estimate the following data:

$$H_x = 3.6 \text{ Bits}$$

$$H_y = 3.6 \text{ Bits}$$

$$H_{xy} \approx 5.3 \text{ Bits}$$

$$T_{xy} \approx 1.9 \text{ Bits}$$

This means that the duration factor introduced by Cohen and collaborators must have contributed about 1 Bit of the 1.15 Bits of partly redundant transmitted information.

The present study is part of a larger programme which aims at the generation of vowel systems of optimal efficiency for the production of artificial speech.

Cohen, Slis, t Hart

Input	Total	Out-put	1	2	3	4	5	6	7	8	9	10	11	12
1	1 670	1 628	42	0	0	0	0	0	0	0	0	0	0	0
2	1 669	82	1 584	0	0	1	0	0	0	0	2	0	0	0
3	1 664	0	25	1 570	0	0	12	0	0	0	57	0	0	0
4	1 670	1	0	47	1 546	68	1	4	1	0	0	2	0	0
5	1 669	4	7	7	3	1 628	0	3	10	6	0	0	1	0
6	1 669	0	0	5	0	1	1 475	128	6	53	0	0	1	0
7	1 670	0	0	0	0	0	126	1 536	2	4	0	2	0	0
8	1 670	0	0	0	0	1	63	10	1 592	0	2	0	2	0
9	1 665	0	30	318	0	0	2	3	0	1 309	1	2	0	0
10	1 663	0	1	221	13	12	19	4	0	28	1 358	7	0	0
11	1 670	0	0	34	274	3	1	5	0	45	322	981	5	0
12	1 670	0	0	0	0	269	1	0	3	0	109	4	1 284	0
Total	20 019	1 715	1 689	2 202	1 836	1 983	1 700	1 693	1 614	1 504	1 792	998	1 293	

$H(X) = 3.58$ $H(Y) = 3.56$ $H(XY) = 4.21$ $T(XY) = 2.93$

Cohen, Slis, t Hart Vowels in one Durationclass Added

Input	Total	Out-put	1	2	3	4	5	6	7	8	9	10	11	12
1	6 680	1 629	1 820	1 545	985	42	71	3	49	5	81	128	322	
2	8 343	86	3	16	6	1 621	1 899	1 605	1 317	1 287	325	66	112	
3	4 996	0	13	132	7	26	13	6	138	1	1 796	1 506	1 358	
Total	20 019	1 715	1 836	1 693	998	1 698	1 983	1 614	1 504	1 293	2 202	1 700	1 792	

$H(X) = 1.55$ $H(Y) = 3.56$ $H(XY) = 3.97$ $T(XY) = 1.13$

Fig. 8.

DISCUSSION

Newel:

The experiments I have performed on human perception of vowel sounds with and without prior knowledge of the speaker would indicate that the remark "as no confusion occurs when we listen to the sounds of a familiar voice" is incorrect with reference to data of the Peterson and Barney type. This will surely invalidate those conclusions made on the basis of this premise.

SOME NOTES ON A NON-ENGLISH SOUND IN CONTINENTAL ENGLISH

MIHAIL BOGDAN*

It is today a phonetic commonplace that "each apperceives the unkown sounds by means of the sounds of his own language" (Boas). When English sounds came into Romanian, however, both in loan words and as English sounds proper in the teaching of the language they came over through the medium of French or German. In these languages the above mentioned rule held good for both the French and the Germans replaced English vowels no. 10 and 11¹, i.e. the vowel sounds in *but* and *bird* by their own front rounded vowel *ö* as in Fr. *deux* and Germ. *schön*. The *Nouveau Petit Larousse Illustré* (1935) records the pronunciation of *trust* as *treust*; *plum-cake* is *pleum-kèke*, while *club* may be either *klub* or *kleub*.

The relationship between the respective English, French, and German sounds has been adequately dealt with by Wilhelm Horn and Martin Lehnert.²

Conversely the French *eu* becomes English *ir* as in *amateur*, while the German names *Röntgen* and *Goethe* become *Runtyen* and *Girte* respectively.

On the more restricted plane of language teaching and learning the same habits prevail and French and German students often pronounce *böt* for *but*, *börd* for *bird*, etc.

The French and German rendering of the vowels under discussion have a justification: a non-French and a non-German sound was replaced by a native one. And since English was first taught in Romania through the medium of French and German these "second-best" vowels were adopted when the Romanian language had within its own vowel system more suitable sounds: *a* for the vowel in *but* and a lengthened variety of *ă* for that of *bird*. It may be noted that the pronunciation of *ö* was, and is, actually difficult for Romanians.

There are two reasons why this vowel was so readily adopted by the Romanian student of English. First, his French and German teachers and text-books taught him so. In the northwestern part of the country the tendency was helped along by Hungarian-English text-books and dictionaries especially as regards the vowel in

* Babeş-Bolyai University, Cluj, Romania.

¹ D. Jones, *An Outline of English Phonetics*, Heffer, Cambridge.

² *Laut und Leben*, Deutscher Verlag der Wissenschaften, Berlin, 1954, pp. 167—184, 454—456.

bird.³ Secondly, the inconsistencies of English spelling may have caused a frame of mind which made the Romanian student look upon English pronunciation as an insurmountable difficulty and the ö-sound being non-Romanian helped the student believe that he was nearer to English the farther he went away from his mother tongue.

A book revealing this frame of mind is *Limba engleză fără profesor, călăuză practică pentru învățarea limbii engleze de toate zilele cu pronunțarea exactă* (English without a teacher, a practical guide to everyday English with the exact pronunciation) by W. Payne and St. G. Gorjan⁴ in the introduction of which the authors speak of the ö-sound (giving Fr. *liqueur* and Germ. *schön* as examples) as an English vowel.

Likewise in L. Duncan's *Gramatica limbii engleze* (English Grammar)⁵ we learn (p. 8) that the phonetic symbol (ə:) represents the long German ö. In his classifications of vowels (p. 12) the author again identifies (ə:) with the long German ö and French *eu*.

The ö-sound was commonly so much associated with English that, at a time, it was thought fashionable to pronounce the name of a brand of cigarettes *Plugar* as *plögär* in a manner considered "à l'anglaise".⁶

It must be noted that English words are frequently pronounced with the ö-sound when the Romanian speaker is especially aware that these words are English. Thus Puccini's opera *Madame Butterfly* is commonly *Madame Böterfly*, *rummy* is *römmy*, *rugby* is often *rögby*. However, some words borrowed from English and which one might expect to be pronounced with an ö, and which have become part and parcel of our language, are given spelling pronunciations and we find the Romanian phonetic value of the English spelling in *trust* and *club* which are pronounced *troost* and *kloob* respectively.⁷

It is true that of recent years, with the spreading of the study of foreign languages, the ö-sound is losing ground and Romanian vowels closer to the English are making more and more headway i.e. *a* and *ă*. The use of plain Romanian *a* in *but* is so much the more recommendable since English vowel No. 10 has moved towards a more advanced and lower position. While some linguists have long marked it in phonetic transcription with a simple (a).⁸

³ Cf. Latzkò Hugo, *Magyar-Angol és Angol-Magyar Kéziszótár*, Budapest, 1929. Dr. Biró Lajos Pál, Dr. Willer Jozséf, Dr. Fest Sándor, *Angol-Magyar és Magyar-Angol iskolai és kéziszótár*. Az Atheneum Kiadása, Budapest. Arthur B. Yoland, *A Dictionary of the Hungarian and English Languages*. First Part (English-Hungarian), Franklin Társulat, Budapest, 1908. Dr. Biró Lajos Pál, *Express English leggyorsabban angolul*, Delta Könyvkiadóállalat, Budapest.

⁴ Editura Gorjan, Bucuresti, 1944.

⁵ Bucuresti, Editura Universul, 1947.

⁶ Iorgu Iordan, *Stilistica limbii romane*, Seria II, Societatea română de lingvistică, Studii, 4, Bucuresti, 1944, p. 39.

⁷ Ida C. Ward, *The Phonetics of English*, Heffer, Cambridge, 1939, p. 201.

⁸ Henry Wyld, *The Historical Study of the Mother Tongue*, London, Murray, 1907, p. 46.

DISCUSSION

Witting:

I think there is a certain parallel in Swedish to what you have told us, since in earlier Swedish pronunciations of the English [ʌ] sound (as in *cut*) the use of [ə] has been quite frequent.

ВОСПРИЯТИЕ ДИФФЕРЕНЦИАЛЬНЫХ ПРИЗНАКОВ И СЛОГОВАЯ СТРУКТУРА РЕЧИ

Л. В. БОНДАРКО*

Определению места слога как наименьшей произносительной единицы речи посвящен ряд глубоких и интересных работ последних десятилетий (1). Можно предположить, что вопрос этот получит определенное конкретное освещение, если принять, что фонема, как минимальная единица языка, реализуется в речи только в слоге. Необходимость установления четких корреляций между дифференциальными признаками фонем и их слоговыми позициями вызвана, в первую очередь, чисто практическими потребностями развития техники связи, для удовлетворения которых оказалось недостаточным описание языка как замкнутой системы, независимой от носителей языка и тех механизмов, которые делают возможным существование языка в его конкретной реализации — речи. Теория дифференциальных признаков — очень серьезный шаг по пути к решению этой проблемы (2). Трудность применения этой теории к практическому (например, автоматическому) анализу речи заключается в том, что инвариантность фонемы как единицы лингвистического кода отнюдь не обуславливает инвариантности физических (артикуляторных и акустических) ее характеристик; а раз это так, то необходимо представить, каковы те механизмы, которые позволяют человеку сводить все многообразие конкретных реализаций к одной единице. Пользуясь только понятием „противопоставление“ — в том смысле, как его употребляют Р. Якобсон и М. Халле, — довольно трудно перейти к описанию этих механизмов. При введении в рассмотрение понятия контраста, роль которого отнюдь не ограничивается выявлением просодических характеристик, задача, как нам кажется, становится более простой. Способ реализации каждого из дифференциальных признаков во всех возможных фонетических позициях зависит не только от числа дифференциальных признаков в системе данного языка, но и от комбинаторики фонем („контраст“).

Рассмотрим с этой точки зрения факты русского языка. Данные анализа объективных признаков и их восприятия убеждают в том, что оптималь-

* ЛЭФ Ленинградского университета — Ленинград.

ными условиями для реализации дифференциальных признаков является слог СГ (3). Примем его за основную единицу при исследовании механизма образования инвариантных признаков фонем. Рассмотрение всех слогов СГ с точки зрения контраста между согласным и гласным показывает, что они организуются в определенную иерархическую систему. При прочих равных условиях (наличие ударения, одинаковый темп) для всех согласных позицией максимального контраста является слог с гласным [a], позицией минимального контраста — слог с [i] или [u]. Остальные гласные занимают пространство между этими крайними точками. Гласные, хотя и в меньшей степени подвергнуты изменениям в пределах слога, тем не менее также в разной степени контрастируют с соседними согласными; в данном случае позиция максимального контраста — соседство с глухим твердым согласным, позиция минимального контраста — соседство с мягким сонантом. Каждый конкретный слог, таким образом, занимает вполне определенное место в данной иерархической системе, при этом место слога определяется как местом гласного, так и местом согласного. Примером слогов с максимальным контрастом являются слоги pa, ta, слогов с минимальным контрастом — слоги vi, li и т. д.

С точки зрения объективных признаков и условий для лучшего восприятия каждого из элементов слога, слоги с максимальной контрастностью характеризуются не только значительными изменениями спектральных (resp. артикуляторных) характеристик во времени, но также и ярко выраженным переходом между первым и вторым элементами, который часто имеет решающее значение для восприятия обоих элементов слога. В слогах с минимальной контрастностью такой переход, как правило, отсутствует. Интересно отметить, что существует определенное противоречие между акустико-артикулярными характеристиками слога и его восприятием; перцептивно оптимальным является слог с наиболее различающимися друг от друга гласным и согласным; с точки зрения акустико-артикулярных характеристик звуков оптимальными как раз являются неконтрастные слоги, в которых признаки гласного и согласного максимально совпадают. Строго говоря, с фонетической точки зрения основной оттенок каждого звука наблюдается только в соседстве с наименее контрастирующим гласным или согласным: для русского [a] это соседство с твердыми заднеязычными, для [e] — с твердыми переднеязычными, для [u] — с твердыми губными, для [i] — с переднеязычными мягкими согласными и т. д.

Ударение выступает в русском языке как один из существеннейших факторов контрастирования элементов слога. Безударные слоги, аналогичные с точки зрения их фонемного состава ударным, отличаются от этих последних ослаблением контрастности, которое является результатом как ослабления некоторых признаков согласного и гласного, так и отсутствием четко выраженной границы между согласным и гласным. В безударных

слогах происходит своеобразное перекрещивание характеристик гласного и согласного, которое в самых ярких случаях (например, в заударных слогах) приводит к тому, что имеется уже не последовательность 2-х единиц, а нерасчленимый звуковой комплекс, несущий информацию о каждом из элементов, представленных в фонологической модели этого слога.

О том, что степень контрастности соседних элементов слога небезразлична для человека при восприятии, свидетельствует целый ряд данных, в том числе ошибочное опознавание звуков.

Необходимо признать, что исследование слога должно занять основное место при описании алгоритмов перехода от многообразия фонетических характеристик к инвариантности дифференциальных признаков. Взаимоотношения между этими двумя системами не являются простыми и однозначными.

ЛИТЕРАТУРА

1. V. Hálá: La syllabe, sa nature, son origine et ses transformations. *Orbis*, X, 1961, N. 1.
2. R. Jakobson and M. Halle, Phonology in relation to phonetics „Manual of phonetics“, edited by L. Kaiser, Amsterdam, 1957, chap. 14.
3. Л. В. Бондарко, Структура слога и характеристики фонем. „Вопросы языкознания“, № 1, 1967.

DISCUSSION

Артемюв:

Я не усматриваю противоречия между понятиями: фонема как пучок дифференциальных признаков и восприятие фонем как совокупность перцепируемых качеств. Человек воспринимает звуковой образ слова, а не сумму фонем.

SIGNIFICANCE OF ELECTROACOUSTIC AIDS IN DEVELOPMENT AND PERCEPTION OF LINGUAL SIGNALS IN THE YOUNG DEAF CHILD

DANUTA BORKOWSKA—GAERTIG*—JANUSZ NOWOSIELSKI*
BOŻENA WIERZCHOWSKA

The development of speech in the deaf child should have a similar course to that of a healthy and normally hearing one, in whom always first appears the period of perceiving auditory notions unrelated to speech, next comes speech-perception, and then only appears the spontaneous speech-production. The method of speech formation in the young deaf child has to be based on the transmission of verbal information about the world it lives in, on rousing interest in it, this method must also tend to develop the child's communicativeness with the world of the hearing and speaking.

In order to afford speech-development to the socially deaf child, it has first of all to be taught to employ the possibilities of its own impaired hearing organ, and this should be done with the use of an amplifier—hearing aid and other acoustic devices.

AUDIOLOGICAL PROCEEDINGS AND AUDITORY REHABILITATION

At the Otolaryngology Clinic of the Institute of Mother and Child, there has been introduced a trial hearing aid for every child with auditory impairment, thus taking account of both children not responding to acoustic stimuli and those responding merely to loud sounds.

Children lacking auditory response proper to their age, are not considered totally deaf, but retarded in auditory development, or socially deaf, thus not supplied with sufficient information by means of hearing.

With patients at the age 1, 2 and 3, responding to acoustic stimuli, the threshold of pain reaction is checked at the initial period, and with older children 2 audiometric curves are drawn initially without, and subsequently with, a hearing aid. The aim is to use binaural hearing aids for the whole day, and for that purpose has been elaborated the scheme (Borkowska-Gaertig) of auditory development. This scheme is divided into 2 periods:

* Institute of Mother and Child (Director: Prof. B. Górnicki, M.D.), Clinic of Pediatric Otolaryngology (Chief: D. Borkowska—Gaertig, M. D.), Scientific Consultant: Prof. I. Cichocka—Szumilin, M.D.

I. — passive i.e. coding auditory notions, unconsciously and consciously. II. — active reproduction of coded auditory concepts.

Within the first period (1) the child is helped to cognize and code, through its subcerebellar and cerebellar centres, human and animal voices, sounds of toys and objects of everyday use, and noises of the immediate environment, (2) next the child is being trained in association processes connected with auditory stimuli, the differentiation of voices and noises from the environment, sounds of various musical instruments, sounds of different rhythms, etc.

II. — After the aforesaid training exercises the reproduction of the coded concepts may be commenced, for inst.: (1) reproduction of rhythm with the use of toys-noisemakers, instruments, a tape recorder, gramophone, and the child's own motoric organs, and its own voice, (2) reproduction of resonance vibration (by sense of touch) of objects emitting sounds (acoustic pressure), instruments, radio, etc., reproduction with its own larynx and "mask" (base) the vibration of the osso-cartilaginous system, (3) reproduction of intensities by modulating the wireless set's dial, tape recorder, amplifying device for auditory rehabilitation, by means of its own voice, (4) reproduction of frequencies with the use of noisemakers, etc., its own voice, and (5) reproduction of sounds from the environment, animal and human voices, such as laughter, crying, shouting and speech.

These preliminary exercises conducted over a period of 1, 2 or 3 years, with regards to the age and developmental possibilities of the particular child, teaches it to differentiate lingual signals from other sounds of the environment, and also prepares for creating spontaneous tendencies of reproducing verbal signals.

SPEECH FORMATION

Speech in the young deaf child is developed with the use of specially elaborated texts (Wierzchowska) for 2, 3 and 4 years old children. The vocabulary of these texts is limited to notions necessary in understanding and in contacts with the environment. For inst., for the 2 and 3 years old have been selected about 100 words, whereas the hearing child uses about 600. These texts are meant for everyday conversations, in the same form as used in contacts with the hearing youngster. Every sentence and each expression is to be used in a relating situation. That is, sentences relating to food are to be used at meals, etc. With children of 2 and older, there may and ought to be used not only actual objects, but also pictures illustrating them, models as well as toys. It is recommended that particular sets of texts ought to be repeated in their original form and as many times as needed for the child to remember and understand them. The criterium for whether the child understands the phrase before it can pronounce it, is the correct response to a message.

ELECTRONIC APPARATUSES FOR SPEECH AND HEARING REHABILITATION

With the purpose of providing for the auditory rehabilitation of children with hearing impairment and lacking speaking skills, there have been constructed (Nowosielski) 2 prototype devices: 1. the transistor amplifier with earphones, 2. the vibration device—an analyzer of the pitch of lingual sounds transposed to sensory vibration.

Ad 1. This amplifying device through microphones receives human voices, music, and other acoustic signals, then after amplification relays them to both earphones. This set is furnished with two microphones: an internal one for self-control, and a second one for the assistant conducting the auditory training. The voice volume control is separate for each earphone in the range of 0—110 dB, that allows of leveling the loudness between both ears. A special output socket and set of small cables provide for connecting it to any wireless set, TV set, or tape recorder. The characteristics of transmitted sounds can be flat in the range of 100—10 000 Hz, or separately controlled for low and high tones. With this amplifier may be used any dynamic earphones, headearphones, or miniature ones with individual ear inserts of a suitable transmitting characteristic in the range of 200 to 6 000 Hz. The resistance of the earphones can be enclosed within the limit of 8 to 60 Hz. This apparatus avails to children and adults with auditory impairment the rehabilitation of hearing and speech, correction of pronunciation defects, cognition of sounds from the environment, and the enjoyment of radio, TV, records, and magnetic tapes.

Ad 2. The vibration unit provides for initiation and correction of speech through sound reception by means of sensing the digital mechanical vibrators. Occurrence of vibrations and their amplitude depend on the volume of high and low frequencies of the tone's components in the voiced sonants of a particular word or syllable. This unit is furnished with a microphone connected to the microphonic amplifier operating with AVC (automatic volume control) compression. The signal of the AVC amplifier's output is transmitted to an 8-channel frequency analyzer containing LC filters. The voltage outputs of the analyzer after detection cause—each in its channel—the power to begin to be amplified and the occurrence of oscillation in the vibrators with a fixed frequency of about 300 Hz. The frequency band for each analyzer's channel, and the appropriate selectivity of the filters, have been selected with particular regard to the phonetic structure properties of the Polish language. The unit is furnished with 8 vibrators, 4 for the left and 4 for the right hand, placed in such a manner that both hands rest similarly to the position on a piano keyboard. The analyzed frequencies increase from the left to the right hand. The aforesaid device enables people with hearing losses ranging on bands of higher frequencies, to differentiate soft consonants such as *ś*, and hard ones like *s* (and also the so called buzzing consonants *s*, *z*, sibilants).

*EFFECT OF USING ELECTRONIC DEVICES IN DEVELOPMENT
OF HEARING AND SPEECH*

After extending the general scheme of consultation for all children with auditory impairment in the country, which has been arranged by the Polish Association for the Deaf, and the Institute of Mother and Child in Warsaw, children from their infancy up to 4 years of age have been the centre of special care. Data compiled in 1967, show that about 75 percent of preschool children with auditory impairment have been supplied with hearing aids, within this number about 45 percent in training to speaking skills, benefit from the assistance of a logopedist, and the rest are subjects to auditory autorehabilitation. The reproduction and understanding of lingual signals in young preschool children have begun to be formed thanks to the earlier development of their auditory perception and cognition of sounds from the surrounding world.

From our initial experience it is deduced, that the early introduction of early acoustic rehabilitation of hearing before the development of speaking skills, forms the foundation for the developmental and intergrational processes of human speech.

THE RELEVANT REGIONS OF CZECH VOWELS FROM THE PERCEPTION POINT OF VIEW

BLANKA BOROVIČKOVÁ*

In the process of determination of the perceptual vowel cues we compare two kinds of experimental material. The first is derived from spectrograms, the second results from listening tests. The linguistic material used was the symmetrical Czech sound groups CVC and VCV. These groups contain all mathematically possible combinations of five Czech vowels and 25 consonants. Each of the 125-member group was uttered by 4 speakers (soprano, alto, tenor, bass).

In this manner we get a total group of one thousand utterances. These groups were recorded on magnetic tape. The record was used first for making spectrograms. We tried to find common or similar signs of the spectra for each phoneme. In our material each vowel phoneme exists in three hundred variants, each consonant phoneme exists in sixty variants. A detailed report on this work was published in 'Spectral Analysis of Czech Sound Combinations'.

The same record was also used for the listening test. The acoustic signal was filtered through half-octave band-pass filters. Fourteen of these filters cover the entire audible frequency range. The signal thus modified was listed by 16 listeners (students of the technical university). The $\frac{3}{4}$ of a million responses to the sounds were evaluated by means of a computer. From this step of our work we got, first the 3000 histograms of the correct responses to each of the sound variant related to the middle frequency band-passes, and second 3000 lists of sound interchanges. Then we divided the responses into three categories of relevancy: Up to 3 correct responses, irrelevant; from 4 to 12, middle relevancy and from 13 up to 16 responses, maximum relevancy.

In the next step we compared both the findings from the spectra and the findings from the listening test to obtain the perceptual cues of vowels and consonants respectively.

I will now try to convey to you our findings, which follow from this comparison of vowels. In addition to the well known facts, which we confirmed by our work, we found some fairly interesting new facts.

The necessity of the presence of two first formants for recognition of the vowel is

* Czechoslovak Academy of Sciences Praha, Czechoslovakia.

a well known assertion. But for the Czech vowel /i/ the presence of one formant only is sufficient, the first or the second formant respectively. Here it is also interesting that the second region of relevancy for /i/ comprises all higher formants (F2 — F5), which makes nearly a single compact region also on the spectrograms. The maximum value of correct responses is in the three bands, with the middle frequency 2,0; 3,0 and 4,5 keps. The middle frequency of the first region of relevancy is 0,2 keps.

Vowel /e/ also has two regions of relevancy, which correspond to both, the first and the second formant frequency. The middle frequency of the first region is 0,4 keps, of the second one 2,0 keps. But the percentage of the correct responses in both regions is very low. This suggests a necessity of the joint presence of at least both first vowel formants.

The third vowel /a/ has only a single region of relevancy at the middle frequency of the band 1,1 keps. It is caused by the fact that both formants of /a/ are close in frequency, only half an octave. So both of the formants fall, at least by their parts, into a single band-pass (1,1 keps) with 100 % of correct responses.

The vowel /o/ has two regions of relevancy (0,6 and 1,1 keps), which merge into a single one. The distance between the first and second formant is again very small, but a little larger than in the case of the preceding vowel /a/. It amounts to approximately one octave. From the fact that the percentage of correct responses is not always in the category of maximum relevancy, we deduce the necessity of the presence of both formants of /o/ for correct recognition.

The last vowel /u/, in spite of the spectral similarity with the preceding /o/, has a quite different characteristic of the relevancy region, developed from the listening test. There is no stable frequency band (region of relevancy) as was found with the preceding vowels. The frequency ranges of maximum of correct responses vary very much and the region of middle relevancy is very large (from the frequency band 0,2 keps up to 2,0 keps). The maximum is in the band of middle frequency 1,4 keps. This maximum is shifted to a higher frequency in connection with the palatal articulated consonants /d̃/, /t̃/, /ň/, /j/ and with the consonants /ž/, /š/, /č/. But according to the spectrograms the band 1,4 keps is neither the frequency of the second nor of the third formant.

Thus we assume, that the main cue for recognition of the vowel /u/ is the transient of its second formant. As the frequency of the first two formant of /u/ is very low, the formant frequencies are determined by a single harmonic frequency of the voice only (particularly by the high voices). But this harmonics varies with the variations of fundamental frequency. This uncertainty of formant determination causes a perceptual orientation towards another cue, the formant transient.

As a conclusion we can say that: these assumptions confirm also our hypothesis about minimal speech unit from the perception point of view. This unit is the consonant-vowel connection and the middle part of this unit, the transients of formants (particularly the second one), is the most important part for both, the consonant and even the vowel recognition.

ACOUSTICAL SUPPORT FOR AFRICAN FIELD WORK

RUTH M. BREND*

This paper will survey some areas of phonological structure of some West African languages which were submitted to spectrographic analysis, and will discuss in some detail the analysis of one phonological problem.

While directing two linguistic workshops in Ghana and in Nigeria¹ Kenneth L. Pike recorded data in which various phonological problems occurred, and sent the recordings to be subject to spectrographic analysis² in the hopes that his on-the-field analyses would thereby be confirmed or refuted. The author of this paper provided Pike with spectrographic measurements and comments thereon.

In general, the spectrographic analysis was of greatest value in the confirmation (or refutation) of the on-the-field, subjective and rather impressionistic justments which had been made, rather than in suggesting completely new avenues of analysis. The measurements of the pitch heights of Izi,³ for example, corresponded exactly, in almost all cases, with the perception of tone heights on the field. The four pitch levels postulated for Igede, similarly, were clearly confirmed by measurements of the corresponding spectrograms, but, for the Bette, Mbembe, and Abua recordings, consistent confirmation of 'down-step' levels was not available. It seems clear that in such cases, the spectrograph is of little value to the linguistic analyst. Although the 'normal' levels of high and low (and, in some cases mid) tones were quite readily confirmed, a complicated tonal morphophonemics would have to be well analyzed, and then special controlled utterances carefully lined up for recording before spectrographic analysis would be useful. On the other hand, the postulated extra-high

* Michigan State University

¹ Under the auspices of the Summer Institute of Linguistics, and partially supported by the U.S. Office of Education, under Contract OE-5-14-065. A summary report of findings of the workshops appears in the report *Tagmemic and Matrix linguistics applied to selected African languages*, by K. L. Pike, U.S. Department of Health, Education, and Welfare, Office of Education, Bureau of Research, November, 1966.

² Charles Peck of the Communication Sciences Laboratory of the University of Michigan made the spectrograms.

³ A language of Nigeria. Of the languages reported on here, Izi, Mbembe, Abua, Igede and Bette are languages of Nigeria; Basare and Twi are languages of Ghana.

pitch syllable of the negative in Abua was easily substantiated by frequency measurements.

In a study of Twi vowels, Pike found that he agreed with J. M. Stewart⁴ in his analysis of tongue-root position being an important feature contrasting the two sets of 'tense and lax' vowels. The spectrograph here was most valuable since, as would be necessary for the hypothesis to stand, a consistent lowering of the first formant for the set of 'tense' (or tongue-root front) vowels was found. (For example, the measurement of the first formant of several utterances of the root-front vowel /e/ was consistently about 200 cps., while the measurements of the first formant of the root-neutral vowel /ɛ/ ranged from 400 to 450 cps.) The implications of this analysis, with its support by the spectrograph, are considerable, not only for Twi, but for other African languages where a similar type of vowel harmony between two sets of vowels has been discovered, but which has been rather impressionistically described previously in vague terms such as 'wide-narrow', 'tense-lax', 'raised-lowered', etc.⁵ From preliminary measurements, it seems that Abua also has two sets of vowels, differentiated by tongue-root position.

Interesting instrumental results were found in the study of Basare rhythm patterns. On the field, Pike and M. Cox postulated a unit occurring in the phonological hierarchy on a level between the syllable and the stress-group (or phonological word) which they called a 'foot'. Feet, within one utterance, seemed to be isochronic, regardless of the number of syllables each foot contained.

They found, for example, that the first feet of the utterances

(1) a ¹ ni	tu ¹ wan	tu ¹ minee	
you know	thing	which-is-here	'you know this thing'
(2) a ¹ caa	tu ¹ wan	tu ¹ minee	
you have	thing	which-is-here	'you have this thing'
(3) a ¹ bati	tu ¹ wan	tu ¹ minee	
you carve up	thing	which-is-here	'you carve up this thing'
(4) a ¹ dakafi	tu ¹ wan	tu ¹ minee	
you taste	thing	which-is-here	'you taste this thing'

that is, [a ni] in utterance (1), [a caa] in utterance (2), [a bati] in utterance (3) and [a dakafi] (actualized in this utterance as [a dauf]) in utterance (4) were perceptually of the same length as were the other feet in the utterance in which each occurred. This analysis was confirmed by spectrographic measurements, with the exception that utterance-final feet of most utterances measured slightly longer than utterance-initial and utterance-medial feet—thus suggesting that utterance-final feet (i.e.,

⁴ See his 'Tongue-Root Position in Akan Vowel Harmony' and Pike's 'Tongue-Root Position in Practical Phonetics' to appear in *Phonetica*, 1967,

⁵ See Stewart, *op. cit.*

feet which occur before pause or silence) are predictably longer than feet occurring in other utterance positions. Measurements of spectrograms of utterances (1) through (4), above, are given below. (An average measurement—in seconds—is given to represent several pronunciations of each utterance.)

a ¹ ni	tu ¹ wan	tu ¹ minee
.32	.34	.4
a ¹ caa	tu ¹ wan	tu ¹ minee
.32	.31	.39
a ¹ bati	tu ¹ wan	tu ¹ minee
.38	.36	.4
a ¹ dakafi	tu ¹ wan	tu ¹ minee
.4	.34	.41

On the field, when making tapes for spectrographic analysis, the informant was asked to tap his spoken utterances. Measurements of the tapped segments, however, did not coincide with the measurements of feet in the spoken utterances. It seems that the informant tapped out in syllables, rather than maintaining the regular foot-timing which was so evident in the spoken utterances. It should be noted, in the spoken utterances, whole syllables disappear in order to maintain the isochrony of the feet (e.g., [a dauf] instead of [*a dakafi] in utterance [4]). This collapsing or disappearance of syllables did not occur when the utterances was tapped—dakafi, for instance, regularly received three taps. The full implications of these differences between the spoken and tapped utterances are not yet clear.

The isochrony of feet also seemed to occur in longer, more complicated utterances, but considerably more analysis needs to be done before the conditioning of feet length can be fully described.⁶

The measurements just described seem to indicate that length in Basare must be treated as a distinctive feature of both feet and individual phonemes (with vowel length marked above as two vowels, etc.). One linear analysis of length as a phoneme would not be sufficient to describe these data. For further discussion of this type of problem see my paper entitled 'Distinctive Features of High-Level Phonological Units'.⁷

⁶ In addition, whenever extra heavy utterance (or sentence) stress had been marked on the field data by Pike and Cox, the spectrographic measurements showed the highest level of pitch (cps.) at that point, thus confirming the prominence of that point of the utterance. In general, also, the foot which occurred with the extra-heavy stress was either the longest in the utterance, or was slightly shorter than the utterance-final foot—thus showing another element in the conditioning of length.

⁷ Paper presented to the Xth International Congress of Linguistics, August, 1967—to appear in proceedings of that Congress.

ЭНЕРГЕТИЧЕСКИЙ КОРРЕЛЯТ СЛОВЕСНОГО УДАРЕНИЯ

Т. А. БРОВЧЕНКО*

Целью данной работы является экспериментальное исследование суммарной энергии ($W_{\text{сум}}$)¹, интенсивности (A) и длительности (t) в создании эффекта ударенности слова в украинском языке.

При подборе экспериментального материала учитывалось, что словесное ударение, являясь одной из характеристик фонетического слова, находит свое проявление в речи на уровне предложения. В экспериментальный материал входили как слова-предложения, так и распространенные предложения различного интонационного рисунка, включающие минимальные пары слов, отличающиеся только местом ударения. Это исключило собственные характеристики гласных и согласных и дало возможность учесть влияние интонационной структуры и позиции слова на акустические параметры словесного ударения. Всего было обработано около 3000 осциллограмм и интонограмм слов и предложений, начитанных 11 дикторами (5 мужчин и 6 женщин), владеющих нормами украинского литературного языка.

Полученные данные показали, что интенсивность и длительность не могут рассматриваться как две самостоятельные акустические характеристики словесного ударения — в украинском словесном ударении они играют только подчиненную роль, как составляющие суммарной энергии.

Главенствующая роль суммарной энергии, как энергетического коррелята украинского словесного ударения подтверждается рядом факторов.

1. В ряде случаев ударный слог характеризовался равным или даже отрицательным значением интенсивности или длительности при положительном значении $W_{\text{сум}}$ и несмотря на это воспринимался всеми аудиторями как ударный.

2. Суммарная энергия в преобладающем большинстве случаев (97,7 %) являлась характеристикой ударного слога в отличие от неударного.

Частота положительных характеристик интенсивности и длительности

* Одесский государственный университет. Лаборатория экспериментальной фонетики

¹ Под суммарной энергией понимается площадь под огибающей интенсивности, измеряемая в мм^2 .

в ударном слоге уступает частоте положительных характеристик суммарной энергии.

3. Процентное отношение положительных и отрицательных характеристик интенсивности (с большей степенью) и длительности (в меньшей степени) ударных слогов варьируется в зависимости от позиционных условий слога, причем эти изменения не остаются стабильными у разных дикторов. Показания суммарной энергии остаются стабильными у разных дикторов и не зависят от позиционных условий слога.

4. Основным фактором является то, что показания суммарной энергии остаются стабильными независимо от интонационного оформления предложения, в то время как составляющие суммарной энергии — интенсивность и длительность обнаруживают большие или меньшие колебания в количестве положительных значений в зависимости от интонационного оформления предложения (фиг. 1).²

Коммуникативный тип предложения	W сум	A	t
Повествовательные	97,8 %	83,1 %	97,5 %
Вопросительные	97,2 %	81,3 %	92,1 %
Перечисления	96,5 %	71,9 %	98,1 %

Фиг. 1. Процентные показатели положительных энергетических характеристик ударных слогов в разных коммуникативных типах предложений.

Обработка процентных показателей положительных энергетических характеристик ударных слогов при различном интонационном оформлении предложений одним из инструментов математической статистики — критерием, определяющим характер расхождения процентных данных показала, что различие между процентными показателями положительных характеристик интенсивности и длительности в ударном слоге является существенным при различном интонационном оформлении предложения, расхождение же в процентных показателях суммарной энергии остаются несущественными.

² Используется формула стандартной ошибки-различия между двумя процентными показателями $SE_{a\%} = PQ \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$, где среднее процентных показателей двух групп $\left(P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} \right)$. См. H. Garrett, Elementary Statistics. New York. London. Toronto, 1956, p. 103.

ными при любой интонационной окраске предложения, что подтверждает превалирующую роль суммарной энергии. Характер вариации указывает на возможное наличие корреляции между интенсивностью и длительностью.

Для проверки гипотезы о наличии зависимости между процентными показателями положительных характеристик интенсивности и длительности в ударном слоге была проведена статистическая обработка данных 48 произнесений слов „руки“ — „рукі“ в словах — предложениях и в нераспространенных предложениях, начитанных шестью дикторами (3 мужчины, 3 женщины).

Статистическая обработка показала наличие линейной зависимости между интенсивностью и длительностью, что подтверждается корреляционными таблицами и коэффициентом корреляции, который во всех случаях показывает высокую степень корреляции (фиг. 2).

		Частота		$\sigma = \sqrt{\frac{\sum x^2}{n}}$		$\sum xy$	$r = \frac{\sum xy}{n\sigma_x\sigma_y}$
		x	y	σ_x	σ_y		
	1-й слог	45	45	4,1	45,6	7148,45	0,85
	2-й слог	48	48	3,8	71,2	8348,88	0,51
	1-й слог (предударные)	48	48	4,5	34,9	5123,01	0,68
	2-й слог (заударные)	42	42	3,6	54,1	6411,90	0,78

Фиг. 2. Коэффициент корреляции интенсивности и длительности (x — количественные характеристики интенсивности; y — количественные характеристики длительности).

Итак, роль интенсивности, длительности и суммарной энергии в создании эффекта словесного ударения в украинском языке не одинакова. Длительность более значима, чем интенсивность, однако как интенсивность, так и длительность играют только подчиненную роль — являются составляющими сложной релевантной характеристики — суммарной энергии ударного гласного или слога, что подтверждается наличием линейной зависимости между интенсивностью и длительностью.

Все вышеизложенное дает возможность утверждать, что суммарная энергия гласного или слога является основным акустическим параметром, отличающим ударный слог от неударного на уровне слова в современном украинском языке.

ОСНОВНЫЕ ТИПЫ ИНТОНАЦИОННЫХ КОНСТРУКЦИЙ В РУССКОМ ЯЗЫКЕ И ИХ ВОСПРИЯТИЕ

Е. А. БРЫЗГУНОВА

В основе краткого изложения интонационной системы русского языка лежат следующие выводы автора:

1. В русском языке есть несколько интонационных единиц, имеющих характерное звучание. Нет синтаксических типов интонации, то есть интонации отдельных предложений. Есть особенности употребления и восприятия одних и тех же интонационных единиц в различной лексико-синтаксической среде.

2. В зависимости от лексико-синтаксической среды интонационные единицы служат средством различения трех уровней значений: а) смысла предложений (например, вопрос-сообщение, виды вопроса), б) смысловых оттенков предложения, в) стилистических и эмоциональных оттенков предложения.

Таким образом, интонационный анализ ведется по двум основным линиям: строение интонационных единиц, употребление и восприятие их в речи.

Путем слухового и инструментального анализа звучащей речи (разговорная речь, чтение текста, дикторская речь) выделены пять основных типов интонационных единиц — интонационных конструкций: ИК-1, ИК-2, ИК-3, ИК-4, ИК-5. Критерии выделения ИК — особенности звучания и способность различать смысл звучащих предложений с многозначным лексико-синтаксическим составом. Строение ИК удобнее всего показать на материале односоставных предложений, в которых тот или иной тип ИК проявляется наиболее ярко:

Она поет. (ИК-1); Как она поет? (ИК-2); Она поет? (ИК-3);

А Наташа? (ИК-4); Как она поет! (ИК-5).

Все типы ИК имеют общие черты строения. Центром ИК является фразовое ударение, которое выделяет составные части: предупредительную-ударную-за-

ударную, предупредительную-ударную, ударную-заударную или только одну ударную часть. Это фонетические позиции ИК:

— ^ — — — — — ^ ^ — — — — ^
Вы бы́ли в кино? Вы бы́ли в кино? Вы бы́ли в кино? Вы?

При инструментальном анализе ИК учитывались длительность, основной тон, интенсивность (общее изменение кривой). При описании ИК-2 употребляется термин усиление словесного ударения (утолщенная линия на схеме). Наиболее существенной характеристикой ИК является изменение тона.

Предударные части ИК произносятся на обычном среднем тоне говорящего, допускающего незначительные повышения и понижения. Характер движения тона незначим для различения смысла и смысловых оттенков предложения, но может служить одним из средств выражения эмоций. На ударной части происходит значимое изменение тона: движение вверх (ИК-3, ИК-5) или вниз (ИК-1, ИК-4) от среднего, нулевое (движение в пределах среднего тона ИК-2). На заударной части тон падает ниже среднего (ИК-1, ИК-2, ИК-3) или повышается выше среднего (ИК-4, ИК-5). Варианты ИК-2а и ИК-4а различаются лишь фонетически: более ярко выраженным повышением тона на ударной части (особенно ИК-4а). Эти варианты широко используются при выражении эмоциональных и стилистических оттенков.

В пределах одного и того же типа ИК возможны вариации по степени повышения или понижения тона, длительности отдельных слогов и слов, усилению словесного ударения. В определенных позициях возникают переходные типы ИК. Эти изменения существенно учитывать при анализе индивидуального произношения.

Как было уже сказано, с помощью ИК различаются три уровня значений. Смыслоразличительными свойствами обладают: а) тип ИК в целом:

Это дом. (ИК-1); Это дом? (ИК-3);

б) изменение синтагматического членения с соответствующей заменой типов ИК:

Этот человек рассказывал нам: /мой отец был хорошим футболистом.

Этот человек, рассказывал нам мой отец/, был хорошим футболистом.

в) передвижение центра ИК:

Никто не открыл ей. И она ушла. (и = тогда)

Он ушел. И она ушла. (и = тоже).

Условие, при котором проявляется смыслоразличительная роль интонации — многозначность (чаще двужначность) лексико-синтаксического состава предложения. Однозначный лексико-синтаксический состав является различительным признаком предложения наряду с интонацией:

Вот мой дом. (слова вот, мой)

Различение смысловых оттенков предложения средствами ИК сводится к следующему: различное осмысление данного и нового, например:

Мне сообщили, что он болен. Мне сообщили, что он болен.

выделение разных членов сопоставления:

Петя едет на Кавказ,^{1/3} а Николай в Крым.¹

Петя/³едет на Кавказ/¹, а Николай в Крым.¹

преобразование пояснения в самостоятельную часть высказывания:

Дом,³ где мы раньше жили,¹/теперь снесли.¹

Дом, где мы раньше жили,³/теперь снесли.¹

изменение смысловых связей между частями сложного предложения:

Он говорил громко,¹ /чтобы все слышали.¹

Он говорил громко,³/чтобы все слышали.¹

ИК могут выражать стилистические и эмоциональные оттенки, если они избыточны для выражения смысла и смысловых оттенков предложения. Так, ИК-3 и ИК-4 выражая в неконечной синтагме повествовательного

предложения незавершенность, различают при этом стилистические оттенки:

Открытие конгресса / в десять часов утра. (разговорная речь)

Открытие конгресса / в десять часов утра. (дикторская речь)

Восприятие стилистических и эмоциональных оттенков предложения в большой степени зависит от взаимодействия интонации и лексики предложения. Так, ИК-5 воспринимается как оценочная интонация, положительное или отрицательное содержание которой конкретизируется лексикой:

Какой ужас! Какая прелесть!

В оценочных предложениях с бытовой лексикой преобладает не ИК-5, а ИК-4а:

Какие туфли я купила!

Таким образом, функции интонации в русском языке (видимо, и в других языках с неслоговой интонацией) определяются природой лексико — синтаксического состава предложения. Многозначность лексико — синтаксического состава, характерная для многих русских предложений, объясняет особенности интонационной системы русского языка.

DISCUSSION

Двонцова:

У меня вопрос, возможно ли считать товарищем Б. приведенные интонационные конструкции самыми основными конструкциями русского языка, которых дальнейшая звуковая модификация их основного (семантического) значения не меняет.

Скоумал:

Каков принцип выделения интонационных конструкций: формальные признаки, значение, или комбинация обоих.

Wiede:

Доклад был очень интересным и полезным, особенно для лиц, изучающих русский язык как иностранный. Очень важно при этом выявление взаимосвязи между лексикой, синтаксисом и интонацией. Возникает вопрос: нельзя ли при исследовании такого рода учесть также и ситуацию, в которой происходит (совершается) речь? Мне кажется, что интонация связана не только с лексико-синтаксической конструкцией, как это правильно утверждает докладчик, но и с ситуацией (средой), в которой совершается речь.

Зиндер:

Доклад, который мы слышали здесь, представляет несомненный интерес, так как автор идет в исследовании русской интонации своими собственными путями. Докладчик отправляется не от синтаксических конструкций, а от формальных мелодических контуров и благодаря этому приходит к интересным результатам. Представленная схема ценна и в теоретическом отношении и при преподавании русского языка.

Брызгунова:

ad Двонцова:

Да, это основные типы конструкций. Их звуковая модификация изменяет не смысл предложений, а лишь его эмоционально — стилистические оттенки.

ad Скоумал.

Учитываются два признака: фонетический и семантический.

ad Wiede:

Ситуация влияет на выбор ик если лексикой — синтаксический состав предложения допускает этот выбор.

Ситуация влияет на выбор звуковых модификаций одного и того же типа ик при передаче эмоционально — стилистических оттенков.

L'INFLUENCE DE LA FATIGUE SUR LA DISCRIMINATION DE LA VOIX FILTRÉE

TEOFILA BYSTRZANOWSKA—BOHDAN DAGAJEW*

Prenant en considération que les lésions de l'encéphale peuvent causer des troubles de discrimination de la voix filtrée (Bocca et col.) nous avons entrepris des recherches sur l'influence de la fatigue sur la discrimination de la parole. La fatigue influence donc entre autres aussi le rendement de l'encéphale.

En 1966 Bystrzanowska a constaté que la fatigue n'influence pas l'audition tonale liminaire ni la discrimination de la parole dans le silence, mais elle diminue sa discrimination dans le bruit (fig. 1).

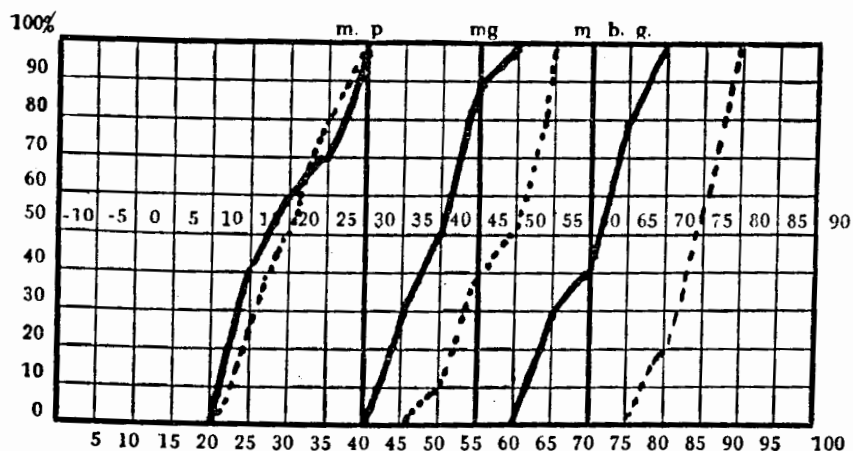


Fig. 1. L'influence de la fatigue sur l'audition dans le bruit: la courbe vocale à l'état du repos; la courbe vocale à l'état de la fatigue.

Dans nos recherches nous avons utilisé maintenant la liste des logotomes monosyllabiques et la liste des mots disyllabiques significatifs, enregistrés sur une bande magnétique et filtrés par 3 types des filtres acoustiques: celui de 63—600 c/s, de 600—1500 c/s et au-dessus de 1500 c/s. Nous avons examiné 10 médecins, 2 fois chacun: après le repos et après 24 heures de service de nuit. Ils étaient presque tous âgés de 30—45 ans. L'examen audiométrique comprenait la courbe tonale aérienne

* I. Clinique d'ORL de l'Institut de Perfectionnement des Médecins, Varsovie.

(d'une oreille) et des courbes vocales de 2 tests non-filtrés ainsi que passés par les 3 filtres ci-mentionnés (16 courbes pour chaque sujet).

Les courbes tonales correspondaient à l'âge du sujet et n'étaient pas altérées par la fatigue (fig. 2).

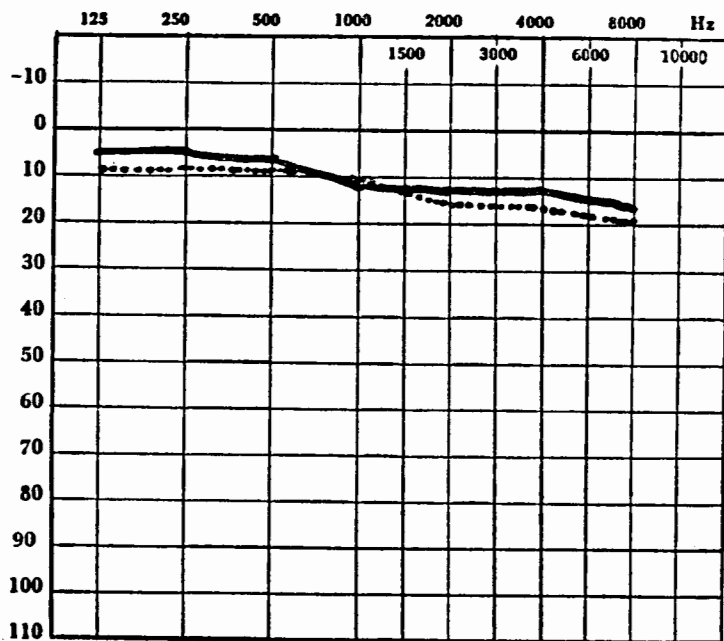


Fig. 2. La courbe tonale: à l'état du repos; à l'état de la fatigue.

L'examen de l'audition de la voix filtrée chez des sujets reposés et fatigués a démontré que la bande acoustique la plus grave (63—600 c/s) est très mal comprise. La discrimination des logotomes dans cette bande ne surpassait d'habitude pas 30%. La discrimination des mots disyllabiques a été un peu meilleure, mais la courbe a été beaucoup plus penchée et rarement a atteint 100% de la discrimination. La bande médiale est un peu mieux comprise, mais même dans cette bande les logotomes ont rarement atteint 100%, la courbe moyenne n'a pas surpassé 50% de la discrimination. La courbe vocale est moins penchée que la courbe vocale de la bande passe-bas, mais elle reste bien différente de la courbe normale et en moyenne elle ne surpasse pas 80% de la discrimination. Dans la bande passe-haut (au-dessus de 1500 c/s) tant la courbe logotomique que la courbe disyllabique sont peu déformées, atteignent presque toujours 100% de la discrimination et leur forme rappelle celle de la courbe non-filtrée (fig. 3).

La comparaison des courbes vocales moyennes chez les sujets reposés et fatigués a donné les résultats suivants:

a) les courbes disyllabiques à l'état de fatigue se trouvent au dessous des courbes des sujets reposés. Cela concerne surtout les tests filtrés (fig. 4). Dans le test non-

filtré l'abaissement touche seulement les faibles intensités- jusqu'à 20 db. Dans les tests filtrés toute la courbe est abaissée, la différence atteint 10—30% et elle est la plus importante dans le test passe-haut qui est le plus facile.

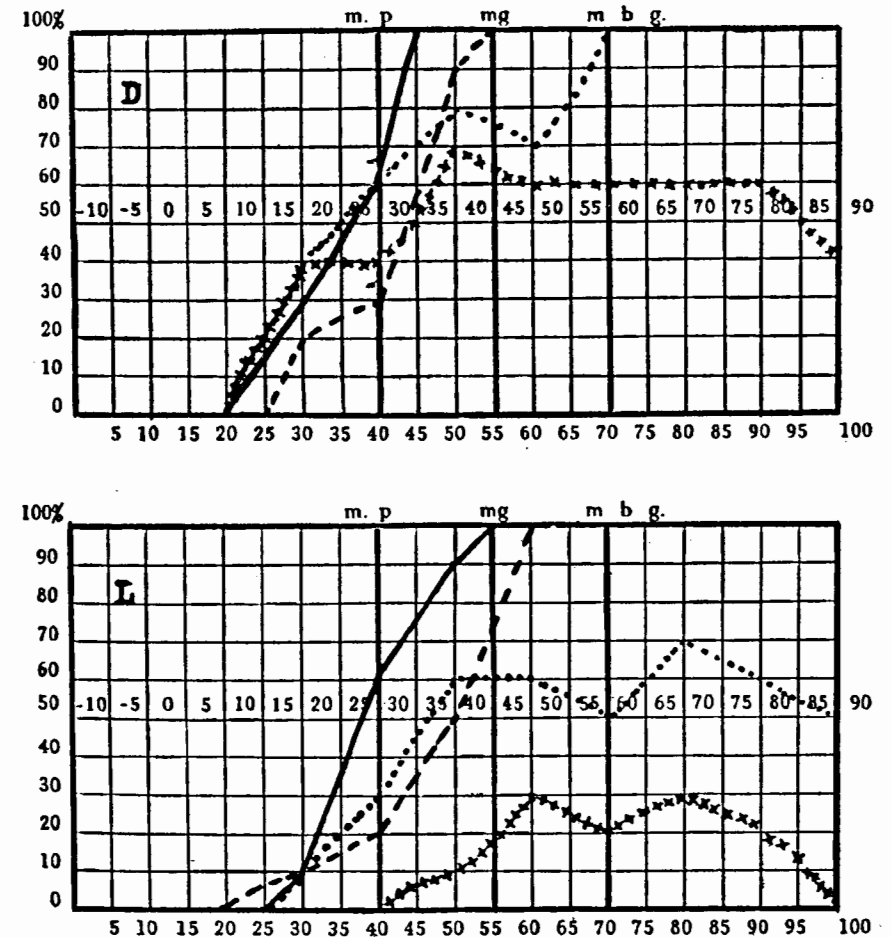


Fig. 3. L'influence de la filtration sur la discrimination des mots: D — mots disyllabiques, L — logotomes: non-filtrés; filtrés 63—600 c/s; filtrés 600—1500 c/s; filtrés au dessus de 1500 c/s.

b) La discrimination des logotomes non-filtrés n'est pas influencée par la fatigue et même la courbe de fatigue est meilleure de quelques pourcents (fig. 5). Les logotomes filtrés passe-bas sont toujours difficiles à répéter indépendamment de la fatigue. Par contre les tests: moins déformé (600—1500 c/s) et très peu déformé (au-dessus de 1500 c/s) sont influencés par la fatigue.

Nous avons considéré les intensités entre 20—70 db comme essentielles: au-dessus de 70 db la courbe vocale pourrait être un peu déformée par le bruit propre de magnétophone. Dans ces limites la différence entre les 2 courbes moyennes était

de 5—20%. Nos recherches confirment le fait bien connu que les fréquences aigües ont beaucoup plus d'importance pour la compréhension du langage que les fréquences graves. L'audition des mots significatifs est beaucoup plus influencée par la défor-

ficatifs disyllabiques les signaux périphériques peuvent être complétés par l'intégration cérébrale.

La fatigue après le travail des médecins présente surtout un caractère d'épuisement nerveux auquel se joint la fatigue physique. Les analyseurs corticaux sont freinés et ils ne perçoivent et ne transmettent plus les stimulus périphériques (Missiuro). Cela s'associe au fait que le nombre des informations transmises par le récepteur périphérique est insuffisant. La fatigue qui chez nos sujets est encore loin de l'épuisement complet, n'a pas encore influencé la perception des signaux primitifs comme des sons purs et n'a pas influencé la discrimination pauvre d'une composition sonore insuffisante (logotomes filtrés passe-bas). Par contre, elle a influencé d'une façon nette l'intégration des tests significatifs, mais déformés. Dans ce cas les centres corticaux n'étaient pas suffisamment effectifs pour intégrer le test vocal qualitativement modifié. Quant au test disyllabique non-filtré, on note une différence seulement dans les faibles intensités. Il paraît que l'information dans ce cas n'a pas été suffisante quant à la quantité: la fatigue a empêché de profiter de ce stimulus trop faible.

On pourrait supposer que la fatigue n'aurait pas dû empêcher de répéter les sons non-significatifs — des logotomes. L'observation montre pourtant que la filtration déforme la composition des consonantes et surtout des voyelles très nettes dans la langue polonaise. La fatigue rend donc leur identification plus difficile et augmente le pourcentage des logotomes méconnus.

CONCLUSIONS

1. Les filtres acoustiques diminuent la discrimination de la parole en dépendance de la hauteur du filtrage: la plus déformée est la bande de 63—600 c/s, la moins déformée — celle au-dessus de 1500 c/s. Donc les sons aigus sont essentiels pour la compréhension du langage.
2. Le filtrage déforme d'une façon plus nette l'audition des logotomes que des mots significatifs.
3. La fatigue diminue l'audition de la voix déformée: cela concerne plus les mots significatifs que les logotomes. Si la déformation est assez importante pour diminuer la discrimination même dans les condition normales, l'influence de la fatigue est moins prononcée.
4. L'influence de la fatigue sur la discrimination des mots nonfiltrés ne peut être notée que dans les faibles intensités.
5. La fatigue semble diminuer la capacité d'intégration de l'encéphale.
6. Les résultats ci-dessus peuvent trouver une application pratique auprès des travailleurs du service de télécommunication.

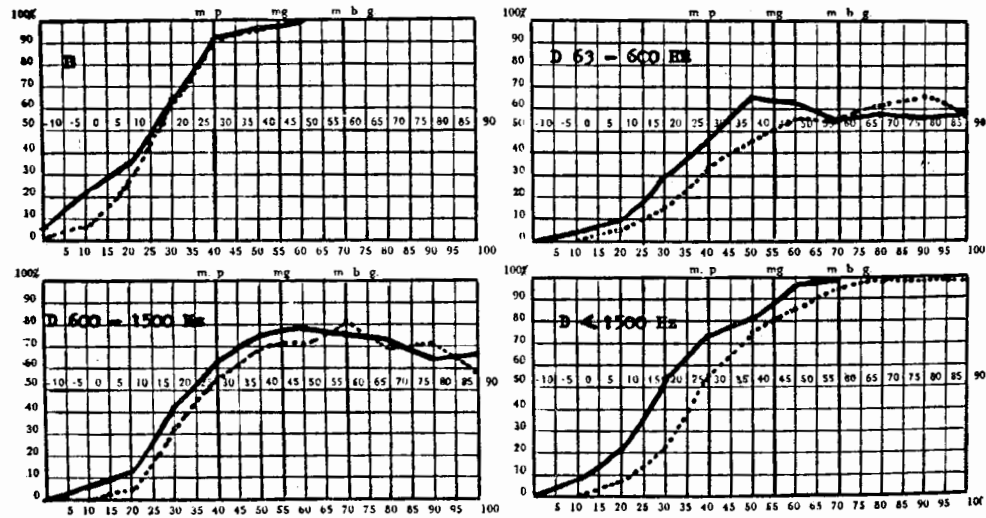


Fig. 4. L'influence de la fatigue sur la discrimination des mots disyllabiques: à l'état du repos; à l'état de la fatigue.

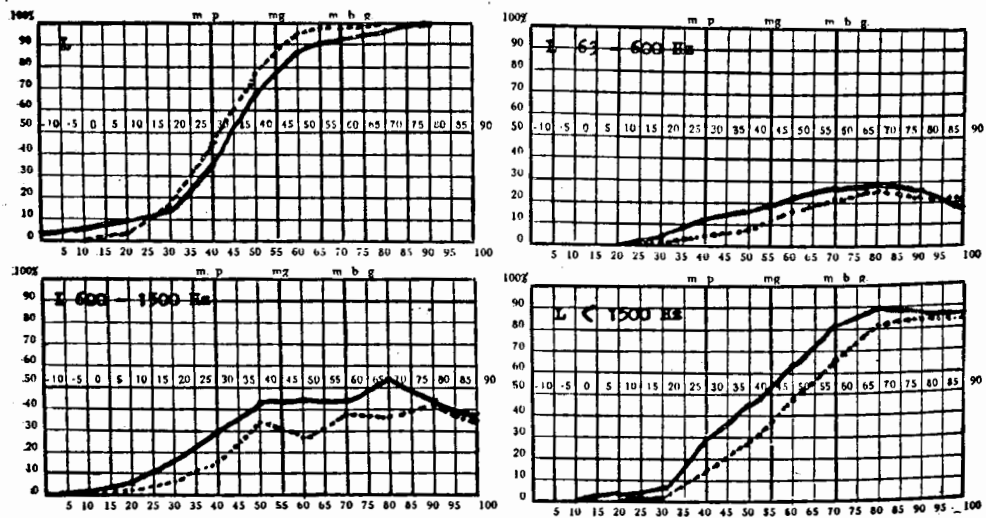


Fig. 5. L'influence de la fatigue sur la discrimination des logotomes: à l'état du repos; à l'état de la fatigue.

mation que celle des logotomes. On pouvait le prévoir car la discrimination des logotomes s'appuie presque exclusivement sur les informations acoustiques livrées par l'analyseur périphérique. Par contre, dans la discrimination des mots signi-

APERÇU GÉNÉRAL SUR LA PERCEPTION DU LANGAGE A REDONDANCE RÉDUITE

C. CALEARO*

Quinze ans se sont désormais écoulés depuis l'observation fortuite — effectuée sur un porteur de néoplasie cérébrale — qui nous a permis d'élaborer dès lors toute une nouvelle méthode d'enquête sur la fonction de la voix auditive centrale.

Auparavant, l'examen audiométrique classique persistait à affirmer, avec une obstination décourageante, l'absence du moindre trouble significatif de la perception auditive en cas de présence certaine — par contre — d'une lésion anatomique du système auditif central.

Chez le malade suivi par nous il y a quinze ans, comme indiqué plus haut, l'ouïe était parfaitement normale pour les tons comme pour les mots, et ce fut avec une réelle stupéfaction que nous pûmes constater au cours de cette enquête un déficit nouveau et particulier de discrimination aussitôt reconnu imputable à l'introduction involontaire, dans le canal de transmission, d'un filtre acoustique qui provoquait la distorsion du message verbal ordinaire.

Cette circonstance accidentelle a été, en quelque sorte, l'acte de naissance d'une sémiologie auditive entièrement nouvelle, dont une série innombrable d'observations expérimentales et cliniques ont par la suite identifié les bases physiologiques, défini les méthodes et contrôlé la validité.

Il a semblé légitime „a posteriori“ que l'identification et l'intégration du message fussent assurées par la multiplicité des éléments neuraux et par l'abondance de leurs relations réciproques même en cas d'inactivation d'une partie des éléments nerveux par suite de la présence d'un processus pathologique; et il a été immédiatement possible de postuler que, dans ces conditions, seule une réduction de la redondance du message auditif pouvait être en mesure de franchir l'obstacle du mécanisme naturel de sécurité et d'identifier enfin un défaut d'opération consistant en une diminution significative de l'intelligibilité.

Dans le cadre de cette nouvelle audiométrie verbale, nous avons opéré la réduction de la redondance de phrases simples ayant une signification autonome: plus précisément, il s'agissait d'énonciations prédictives, selon un modèle - base constitué par le schéma sujet-verbe-complément.

* Clinique Universitaire d'Oto-rhino-laryngologie de Ferrara, Italie (Directeur: Prof. E. Bocca).

A défaut de tabulation de la fréquence des termes dans la langue italienne, les mots des phrases en question ont été choisis de façon à utiliser uniquement ceux que l'usage commun emploie le plus fréquemment. La longueur des phrases allait d'un minimum de 8 syllabes à un maximum de 16, avec un mode de 10, ce qui donne une idée de leur durée.

Partis du mot, nous avons choisi par la suite la phrase comme matière définitive d'examen. La phrase constitue en effet, grâce à ses liaisons non seulement phonétiques et grammaticales, mais aussi syntactiques et sémantiques, quelque chose d'extrêmement redondant et en même temps d'élastique et de maniable, où la redondance peut être réduite de façon plus variée qu'on ne le pourrait par l'emploi des mots purs et simples.

On peut suivre à cet effet différentes modalités. En ce qui nous concerne, nous avons préféré opérer la sensibilisation du matériel verbal au moyen de deux systèmes assurant une distorsion différente: le premier mis en oeuvre sur le spectre du message (voix déformée), le second sur le temps (voix accélérée).

Dans notre *matériel accéléré*, un système particulier de réduction du temps élève la vitesse d'émission des mots de la valeur de 150 mots/minute (normale pour la langue italienne) à celle de 350 mots/minute.

Le système employé se base sur l'usage de deux magnétophones munis d'un dispositif de variation de la durée du temps d'enregistrement, sans altération simultanée de la fréquence des sons enregistrés. Dans ce dispositif, la tête d'enregistrement est en mouvement par rapport à la bande magnétique, de sorte que la vitesse relative demeure constante.

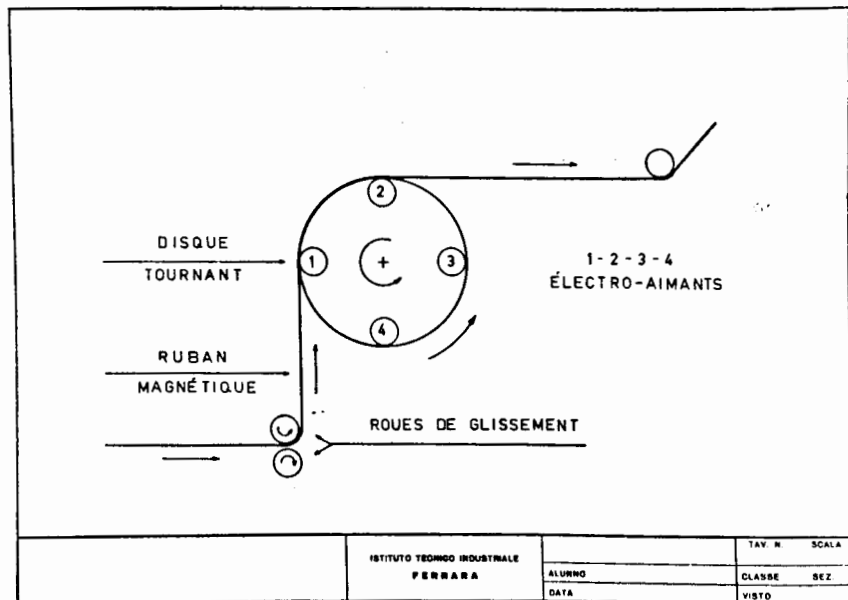


Fig. 1.

Pour atteindre ce résultat, la bande est reproduite par un groupe de têtes en mouvement (comme l'indique la figure) et non par une tête fixe.

2. Dans notre *matériel déformé*, un procédé particulier de filtrage et de distorsion réduit la redondance.

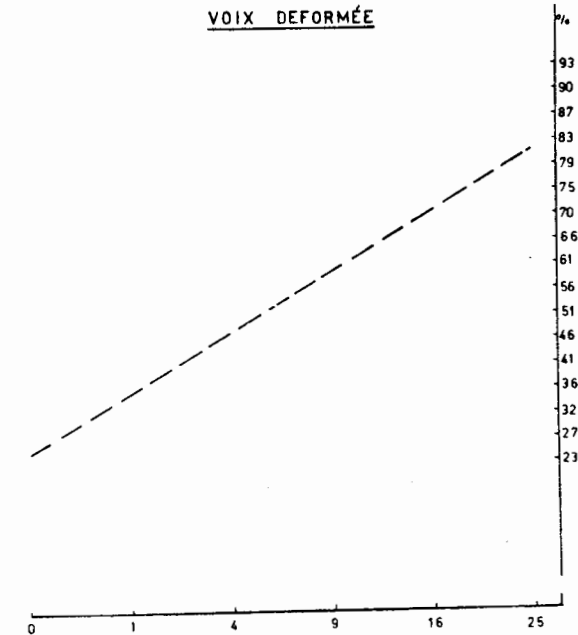


Fig. 2.

Le filtrage est obtenu en soumettant les bandes de fréquence 400—800, 1600 - 3200, 3200—6400 Hz à une atténuation progressivement croissante. On atteint ainsi un effet d'ensemble comparable à un filtre passe-basse 500 Hz. La distorsion est réalisée en faisant passer le matériel filtré par un ring-modulator, construit selon le système classique et destiné à produire les tons de combinaison.

En ce qui concerne l'évaluation des résultats, l'intelligibilité a été mesurée de la façon suivante: nous avons considéré comme correctes les répétitions sans erreurs ainsi que celles qui étaient entachées de quelques erreurs de détail (il s'agissait d'ordinaire d'un singulier à la place d'un pluriel; ex.: „l'enfant joue dans les prés“ au lieu de „les enfants jouent dans les prés“).

Ces deux types de réponse correspondent respectivement au 1^{er} et au II^e type de la classification de Chavasse et excluent les fautes de grammaire, les aberrations sémantiques, les réponses ne correspondant pas à l'original transmis, enfin l'absence de réponse.

Chez le sujet normal, la valeur maximale d'intelligibilité atteint le 80 % à un niveau commode d'audibilité, pour la voix accélérée comme pour la voix déformée.

La standardisation des tests — autrement dit la détermination des champs de

variabilité de l'intelligibilité chez le sujet normal — revêt une importance considérable aux fins cliniques.

Nous avons effectué à ces fins des recherches très considérables au cours desquelles l'intelligibilité du matériel vocal sensible a été considérée comme une variable

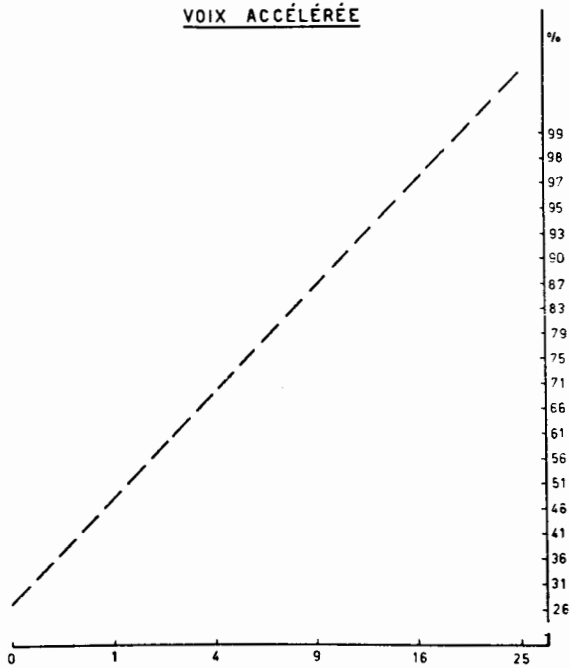


Fig. 3.

dépendant de plusieurs facteurs. Nous avons pu constater que deux seuls des facteurs en question influent sensiblement sur l'intelligibilité: l'âge du sujet et la scolarité.

Nos deux figures représentent l'intelligibilité (en ordonnées) selon le nombre d'années d'école (en abscisses) chez 100 individus normaux d'âge compris entre 20 et 30 ans.

La première figure se rapporte à la voix déformée, la seconde à la voix accélérée. Nous voyons que l'intelligibilité augmente en fonction linéaire de la scolarité. La zone entourant les courbes représente les limites de confiance à 95 %.

Nous disposons ainsi d'un instrument précieux pour évaluer le caractère plus ou moins normal du résultat obtenu chez les malades examinées.

A NOTE ON THE VOWELS OF BACHAMA

J. CARNOCHAN*

Bachama is one of the Bata dialects, spoken by some 50,000 people in the Adamawa area of Nigeria, and in the adjacent territory of the Cameroons Republic. Like Hausa, it is classified as Chadic. Very little has been written on the language, but several scholars making preliminary studies have spoken of difficulty with the vowels. I found this too, particularly with two unrounded back to central vowels, transcribed here *u* for the close and *ø* for the half-close quality. I would hear the same word one day with one quality and the next day with the other, and doubted the accuracy of my perception until I found that indeed both qualities are regularly used in certain items, but not in all, according to the grammar of the 'piece'. I am illustrating this characteristic of Bachama here with examples of short verbal sentences, showing the different sorts of concord found.

1. *nda ðim ko:mboto*, -\ _ _ - - He sank the canie.
2. *nda ðyem ko:mbye:*, -\ _ _ - - He sank the canoes.

In 1 there is concord between the singular form of the verb *ðim* and the singular form of the noun object, *ko:mboto*; and in 2 between the plural forms *ðyem* and *ko:mbye:*, the differences between the two verb forms including different vowels. Both singular and plural forms are palatalized, this feature being emphasized in the plural form by the letter *y* in the transcription. There is a similar singular/plural concord between verb and object in the Aessive aspect forms, e.g.

3. *nda ðima: ko:mboto*, --\ _ _ - - He went and sank the canoe.
4. *nda ðyama: ko:mbye:*, --\ _ _ - - He went and sank the canoes.

A second type of concord is found between the subject and the verb, as in the following.

5. *ko:mboto a ðimo*, _ _ _ _ - - The canoe sank.
6. *ko:mbye: a ðyemo*, _ _ _ _ - - The canoes sank.
7. *nda zumo*, _ _ _ He ate.
8. *ta: zomo*, _ _ _ They ate.

* School of Oriental and African Studies, University of London.

9. *ŋga 'uso*, -⁻-₋ She cooked.
 10. *ta: 'oso*, -⁻-₋ They cooked.

These examples have short vowels in the verb stems, but there are similar features with long vowels, e.g.

11. *nda hi:wo*, -⁻-₋ He prayed.
 12. *ta: hye:wo*, -⁻-₋ They prayed.
 13. *nda tu:lo*, -/₋ He chewed.
 14. *ta: t:lo*, -/₋ They chewed.

So far no verbs have been found with long -*u*. Verbs with a long vowel when the syllable is open, have a short vowel when it is closed, e.g.

15. *nda pi:ro*, -⁻-₋ He thatched.
 16. *nda pir vane*, -⁻-₋ He thatched the hut.
 17. *nda pyer vonye*, -⁻-₋ He thatched the huts.

Sentences 16 and 17 show concords of the first type, with the singular and plural forms of the noun object.

In addition to the two types of concord illustrated above, there are pairs of sentences with differences in the verb stem vowels of the sort associated with the singular/plural terms in the category of number, but where there are no concords with other elements in the sentence. E.g.

18. *nda mbura diye*, -⁻-₋ He went and put out the fire.
 19. *nda mbara diye*, -⁻-₋ He went and beat out the fire.

Sentence 19 draws attention to the number of acts performed in extinguishing the fire, while sentence 18 does not. The other items in the sentences are the same for both. Here are examples where the relation of the singular or plural verb form is with the situation, and not with any other item in the sentence.

DISCUSSION

Dunstan:

Is there a one-to-one correspondence between the vowels of each pair of concord sets?

Goodman:

Might the phenomenon described correlate with separateness versus unity of the action rather than with singularity versus plurality of the object or subject.

Carnochan:

ad Dunstan: No, there is not a one-to-one correspondence. We have -*o* in 9 and 10, but there is also -*u* in *nda tuwo* he wept; *ta: tɔwo*, they wept; and -*u* in *nda puwo*; *ta: powo*, they bent down. The concord is not just a matter of vowel relationship, but requires other features of the syllable as a whole to be taken into account.

ad Goodman: The evidence offered by the last pair of examples points in the direction suggested by Mr. Goodman, and all the other examples could be treated in this way, which would be by reference to the situation. It nevertheless remains that a large number of examples show concords of the types mentioned, and in a formal analysis, one would wish to draw attention to them where possible.

PENTE ET RUPTURE MÉLODIQUE EN FRANÇAIS RÉGIONAL DU NORD

Analyse et interprétation d'un trait prosodique caractéristique

FERNAND CARTON

Dans de nombreuses régions, en français populaire ou en patois, la pénultième de mot ou de groupe rythmique s'allonge souvent: „Dans le récit, un rythme instinctif s'établit; la quantité affecte la syllabe précédant la tonique“.¹ En français cultivé, les variations de fréquence du Fo peuvent atteindre une octave, mais on ne ralentit guère, alors qu'en parisien populaire, on prononce: „A Pâris“, en Lorraine: „Nâncy“. de même à Liège, en Savoie² etc. Mais, à y regarder de près, ce trait est bien différent selon les parlers. Il nous semble que les parlers populaires picards de la région Lille-Tournai se prêtent bien à une étude de ce genre: auditivement, nos informants et nous-même nous percevons souvent un rythme trochaïque, avec „chute“ sur la pénultième, la „brève“ étant plus haute. Tous nos informants voient là un recul de l'accent; d'autre part la présence d'enclitiques, que l'on peut accumuler avec un ton „uni“, ne modifient pas le trait. Un e caduc maintenu s'allonge et „glisse“ comme toute autre voyelle. Ex.: *Quo qu'te fais, ti?* „Que fais-tu, toi?“

Méthodes d'enquête: 1. Sondage, dans le cadre d'une enquête dialectologique aux fins diverses, dans l'arrondissement de Lille, dans des milieux socio-culturels divers. Originaire de cette région, nous avons pu noter au vol, dans des conditions de spontanéité excellentes, des nuances précieuses. — 2. notation et enregistrement de conversations et de récits „spontanés“, au hasard des occasions, dans un plus large secteur. La technique du „faux arrêt“ permet de bonnes observations: durant les pauses d'un questionnaire-prétexte, on continue l'enregistrement tandis que les informants se détendent. — 3. enregistrement de phrases isolées avec et sans emphase présentant des voyelles semblables aux 2 dernières syllabes pour faciliter la comparaison syntagmatique à l'instrumentation — 4. écoute par des informants des enregistrements 2 et 3 en demandant s'ils reconnaissent comme locaux les traits prosodiques qui y apparaissent. — 5. des fragments de chaîne parlée s'avérant avoir un sens différent selon qu'ils sont ou non pourvus du trait, on demande pour chacun le sens précis. — On a interrogé à chaque point 2 témoins d'âge nettement différent; un point par tranche de 5 000 hab. en ville, et 1 pour 3 000 ou moins dans les villages. (Ecoute d'enregistrement.)

¹ Abbé Rousselot, *Revue de Phonétique*, I, 1. p. 89.

² Martinet, *La description phonologique (Hauteville)*.

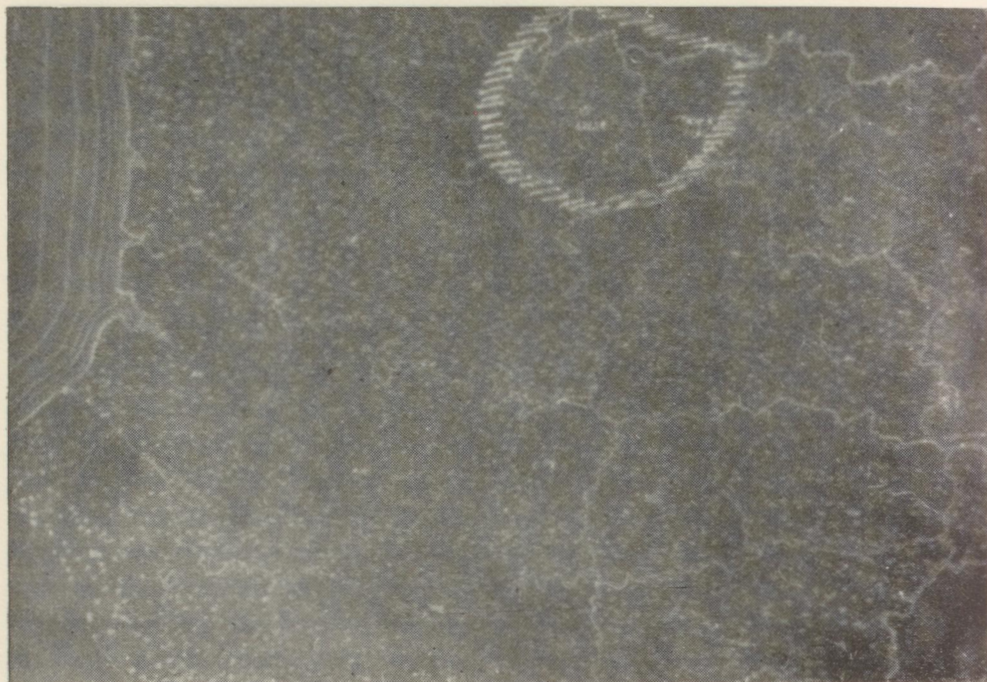


Fig. 1.

Zone: Ce sont les méthodes 2 et 5 qui ont semblé les plus sûres par recoupement. La 4^e n'a servi que pour l'instrumentation pour les points où le trait a été reconnu par les méthodes 2 et 5. Des comptages ont permis de tracer une „limite“ (fig. 1) de fréquence „forte“ (plus de 60 % compte tenu des déterminations ci-après). Un pourcentage supérieur à l'intérieur de cette zone se remarque dans la conurbation urbaine Lille-Roubaix-Tourcoing.³ Cette zone est minimale, puisqu'on entend ce trait au-delà, mais la fréquence y est moindre. Elle recouvre en gros l'ancienne Châtellenie de Lille et l'ancien évêché de Tournai et recoupe plusieurs isoglosses morphologiques et phonétiques. C'est dans les milieux populaires qu'il connaît la plus grande extension: la détermination géographique doit donc se doubler d'une détermination *socio-culturelle*. Du point de vue de l'âge, nous ne pouvons guère conclure actuellement: la fréquence de manifestation est plus grande chez les personnes de plus de 40 ans, mais n'est-ce pas parce que celles-ci parlent davantage le patois? Patois et français régional sont étroitement imbriqués; ils semblent avoir une prosodie identique. Enfin on n'a tenu compte que de faits relatifs à une même *situation*, une surcharge affective rendant la réalisation, sur le plan de la substance, bien mieux décelable.

³ Le trait est également présent parfois chez des ouvriers flamands frontaliers, mais cela pose des problèmes que nous n'avons pas encore étudiés.

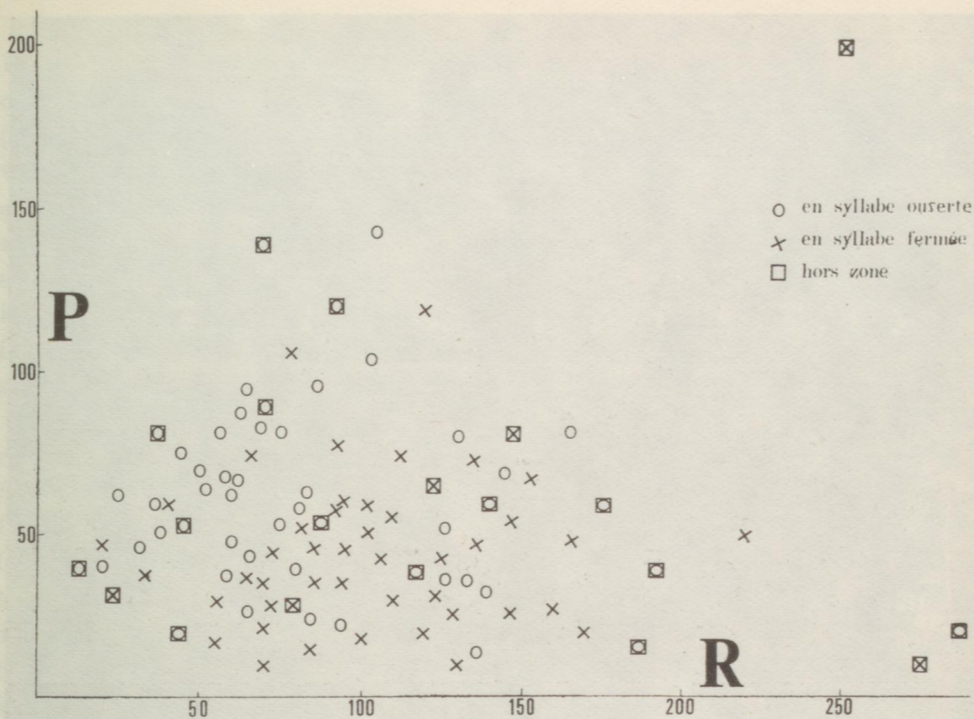


Fig. 2.

Etude instrumentale des corrélats physiques:⁴ Les divers paramètres de l'accentuation obtenus, nous avons réalisé des graphes portant sur 100 cas. Le changement de „direction“ du ton des 2 dernières voyelles de groupe, la pente et la rupture mélodiques nous ont paru caractéristiques (fig. 2, 3). Par *pente* P (point de vue instrumental) nous entendons la baisse de fréquence du Fo combinée avec la durée (accompagnés d'une baisse du niveau de pression sonore). Une chute mélodique (point de vue auditif) de 5 quarts de ton en 10 cs est mieux perçue de nos informants qu'une chute de 5 qt en 5 cs, moins étalée dans le temps.⁵ — Dans le premier cas le coefficient intervalle-temps est 50, dans le 2^e de 25 —. Pour une pénultième en syllabe fermée par une occlusive, la pente est raide et donne une impression d'inachèvement, la durée de tenue de cette occlusive est accrue; ce n'est pas la consonne initiale de la syllabe suivante qui s'allonge. Ex: c't'aut' chose“ c'est autre chose“. Des assimila-

⁴ Faite au Laboratoire de Phonétique de NANCY avec l'indicateur de fréquence et l'indicateur d'intensité construit par Frøkjær—Jensen et inscription par oscillomink. On a vérifié en comptant les phases, et fait également des spectrogrammes pour quelques fragments difficiles. Par référence avec les lignes d'intensité, nous avons négligé le début et la fin, mal stabilisées, des courbes mélodiques des voyelles.

⁵ Nous avons préféré ce produit mieux en rapport avec l'effet auditif, au coefficient de direction \underline{a} de la fonction linéaire $y = ax$.

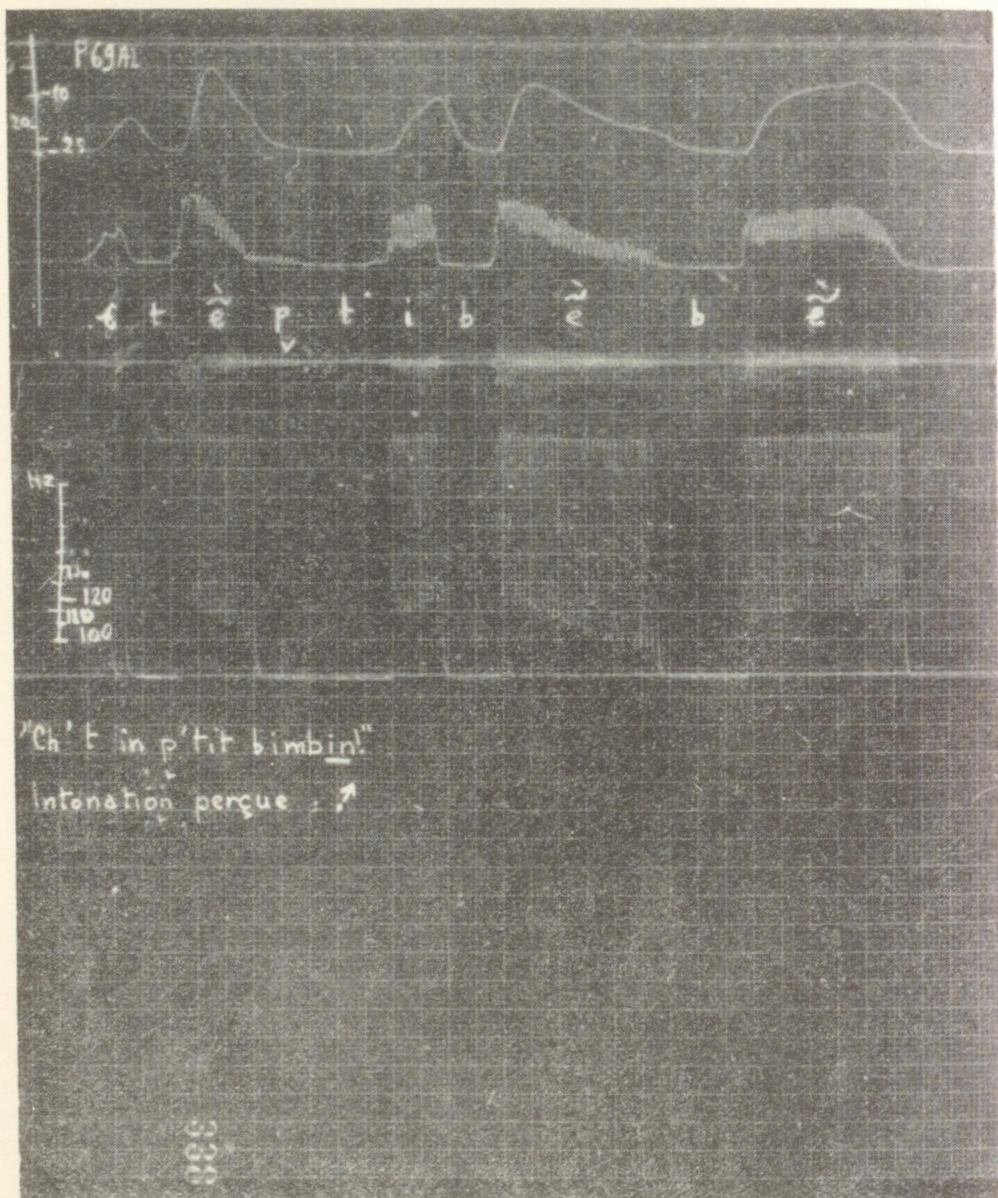


Fig. 3.

tions diverses (*tout l'temps* > *tout t'temps*, *tchubute* > *tchubbute*, etc.) paraissent favoriser la „chute mélodique“ de la pénultième et cet allongement consonantique. Nous avons jugé intéressant de mesurer, en faisant également intervenir la durée, la *rupture mélodique*. R⁶ Interruption consonantique de 5 cs combinée avec

⁶ Déjà étudiée pour le français par Coustenoble-Armstrong, G. Faure, etc.

intervalle de 10 qt: coefficient intervalle-temps de rupture: $R = 50$; interruption de 10 cs et intervalle 10 qt : $R = 100$. L'effet auditif du trait est accru par l'importance de la discontinuité. Il y a comme une économie d'énergie: la chute de la pénultième (compte non tenu des enclitiques) paraît permettre de monter moins sur la dernière voyelle, qui a cependant un grand relief grâce à la rupture. C'est la graphe P/R qui a semblé le plus intéressant. Le groupage des points montre entre autres que ce trait se réalise plus par R que par P. — Nous avons séparé les cas où la voyelle pénultième est en syllabe fermée: P y est moindre, pour un effet auditif à peu près semblable. D'autre part, ce graphe peut fournir une confirmation pour le tracé de la zone de fréquence du phénomène étudié: sur 20 points qui n'ont pas été relevés dans cette zone, 13 se situent en marge.

Interprétation linguistique: 1) Tout se passe comme si la pente avait pour rôle de préparer la mise en relief de la dernière voyelle. Nous sommes en présence d'un *fait d'accent* de groupe rythmique, réalisé par la *combinaison prétonique + tonique*. Une diphtongaison peut affecter la prétonique (Ex: *léongs ch'veux*) „longs cheveux“. Comme il n'y a jamais d'autre accent dans un groupe quand ce trait est présent, c'est à la fois un accent expressif et un accent rythmique: plus précisément l'accent démarcatif de groupe se renforce d'émotivité. Le trait existe aussi dans les interrogatives, bien que P et R y soient plus faibles. — 2) Ce fait d'accent se superpose à un *fait d'intonation*: c'est particulièrement sensible dans les couples prosodiques dont nous avons parlé, où le trait prosodique confère à l'énoncé un sens qui n'est compris que des natifs (Ex.: *j'comprends* = „c'est certain“).

Ce n'est plus un emploi passif et mécanique, mais une modification intentionnelle. L'actualisation diffère: l'achèvement (un ou plusieurs sens) s'oppose à l'inachèvement, avec valeur „exclamative“ et sens particulier au pays.

Conclusion: Tracer des isoglosses prosodiques nécessite beaucoup de précautions, une étude instrumentale complémentaire, et l'observation de *petites aires*, aux parlars relativement proches les uns des autres. Ainsi l'étude de la prosodie des français régionaux, aux traits fortement marqués, contribue-t-elle à l'étude de l'accentuation du français standard, qu'on peut définir négativement. Enfin ces recherches montrent l'importance que présente la combinaison durée + variations de fréquence.

DISCUSSION

Cohen:

L'étude des quantités liées à l'intonation doit évidemment être faite de manière à essayer de délimiter des zones. Certains parlars de régions et de milieux ont des quantités prolongées de syllabes finales.

ACOUSTIC ASPECTS OF POETRY

P. P. J. VAN CASPEL

De Groot (1946, 48) once made a remark about what he called one-sided positivism shown by many scholars representing experimental phonetics, who neglect the existence of a psychic background to actual phenomena. Elsewhere (1957, 399), he admitted that it seemed without doubt that the solution of several problems as to structure and recitation of poetry could only be expected from the study of the sound waves and the body movements concomitant with the production of these sound waves. The sound form of a poem, he held, comprises linguistic features only. In his opinion, recitation as such is a matter of actual sounds, not of sound form (1957, 389). In an earlier study, though, he has shown that one cannot always draw a clear dividing line between verse structure and verse recitation (1946, 80). The question may be raised, however, whether it is not the actual recitation which gives rise to the verse structure as such. Above all in those cases in which the poet recites his own work, it might be said that realization of the sound form and poetical structure are identical. This might be true from the listener's point of view as well. As L. Kaiser says: With the help of somatic factors as go-betweens, psychic factors make their appearance in acoustic structures. The ear in its turn makes it possible for these acoustic structures to create psychic phenomena in the hearer's mind (1953, 96). Now, that which is experienced by the hearer as a verse structure, is such a psychic phenomenon in its own right. In order to throw some light upon this elusive phenomenon, one might tackle those underlying acoustic structures, using the simple methods of experimental phonetics. I fail to see why this should be branded as one-sided positivism. — In the Laboratory of Experimental Phonetics of the University of Amsterdam, six young poets were asked to read some poetry of their own as well as some unpublished work by others. Moreover, they were subjected to some tests (counting till ten, Stanford, talking about pictures) which, in previous experiments, had been applied to groups of students and school children. The Stanford score resulted in an 'originality index' of 80 for the poets, 73 for the children, and 52 for the students. In the spacing of the words within the allotted period of 60 seconds, too, poets and children showed a similar picture, clearly different from that produced by the students. Finally, the poets' respiratory movements were recorded by means of a pneumograph. Differences in structure and rhythmical flow of certain poems could be traced back to striking differences in the respiratory curves of their

respective authors. --- The results of the recordings made by means of tapes and gramophone records, were then analyzed as to duration, stress, and pitch. As far as possible, the values found by means of kymograph or oscillograph recordings, were visualized by transforming them into graphs. In most cases, some standard of time was recorded at the same time in order to have a measure of duration. — As to duration, i.e. the time relation between sound unit and pause, it was found that the picture presented by the printed text (the lay-out of the poem) did not always conform to the acoustic reality. In trying to establish, as De Groot seemed intent upon doing, the correspondences between successive verses or lines of a poem, one should, therefore, consider first of all the exact definition of a verse or line. — Stress is rather hard to define. That which the listener experiences as speech intensity, as accentuation of individual syllables, does not, primarily, depend upon the absolute intensity of sound, the quantity of which can be measured by objective methods. It is rather the distribution of certain doses of energy, however weak in themselves, that gives the impression of accentuation. Moreover, an increase in energy may at the same time entail an increase in duration and pitch. It was found — a 2-syllabic word, in the same context, being spoken by 7 subjects—that the hearer's subjective impression of accentuation or stress was not directly related to the sound intensity recorded by objective methods. In the same way, looking at graphs representing both stress and pitch curves of certain lines or verses, combined with listening to the sound recordings involved, made it clear that accentuation did not always depend upon the absolute sound intensity of the syllables constituting the line. Much more important was in each case the relation to the surrounding syllables. In some cases a decrease of intensity was compensated by rising pitch and prolonged duration in such a way as to maintain nonetheless an impression of accentuation.—The syllable pitch series, forming the verse melody, did not always show a particularly striking profile. In many cases the melody curve remained within the boundaries of normal speech. Aesthetically relevant, however, proved to be the presence in the line—even within rather flat melody curves—of at least two rises in pitch, spaced at a certain distance. This was especially effective when such a structure was repeated in the next line or the next half-verse. Clearly, it is this phenomenon that is of fundamental significance for the rhythmical experience in poetry. The results of the experiments described might lead us to believe that it is the melody rather than the sound intensity which produces the nuclei of intensity from which the hearer derives his experience of rhythm.

REFERENCES

- A. W. de Groot, *Algemene Versleer*. Den Haag 1946.
A. W. de Groot, Phonetics in its relation to aesthetics, in:
Manual of Phonetics, ed. L. Kaiser. Amsterdam 1957.
L. Kaiser, *Les sons du langage et leurs informations*. La Voix, Paris 1953.

DISCUSSION

Hannafoord:

Many previous studies of poetic performance are open to the objection that they were made on very simple poetry. In these studies the most important question of how competence is related to performance is, in fact, avoided. What procedures, if any, did you follow in your experiments to test the degree to which the readers of poetry in your study 1) were performing poetry of some difficulty and 2) understood the semantic and syntactic structures of the poems they were reading?

van Caspel: *ad Hannafoord*

The poetry recorded was of a highly complex and even controversial nature, the authors belonging as they did to the post-war school of so-called experimentalists. The poets were asked to read their own poems and their performance was, on the whole, utterly convincing. Apart from this subjective impression, it was found that the pitch patterns of the individual lines were, mostly, of the convex type, which is known to convey a speaker's inner certainty.

ИСПОЛЬЗОВАНИЕ СВОБОДНОЙ РЕЧИ В ФОНЕТИЧЕСКИХ ИССЛЕДОВАНИЯХ

L. CERLITIS

Об использовании фактов т. н. свободной речи для экспериментально-фонетических исследований говорится уже давно. Однако число работ, основанных на таких фактах, все же остается небольшим. Тем не менее поток свободной речи может давать интересный материал для акустического анализа речевого сигнала, особенно изменчивости, нестационарности его элементов (звуков, слогов, интонации и др.). О такой нестационарности не всегда свидетельствует дикторская речь, т. е. факты, продуцированные специально для данной серии измерений. Диктор фактически говорит в обстановке, близкой к ситуации публичной речи. Если он не в состоянии преодолеть всем известные факторы, тормозящие процесс речи при такой обстановке, звучание речи становится более стабильным, чем в повседневной обстановке.

Разумеется, я отнюдь не отрицаю значения специально продуцированной речи для экспериментальной фонетики. Здесь будет говориться только о расширении объекта и о понятиях, необходимых, на наш взгляд, при оценке качества выбранного для анализа материала.

В первую очередь приходится уточнить понятие „свободная речь“. Как известно, в литературе свободной считается или 1. импровизированная речь, текст которой образуется во время выступления (в противоположность репродуцированной речи, или 2. речь, которая произносится наизусть, или 3. хорошая речь („свободно говорить по-немецки“ фактически означает — хорошо говорить). Я буду использовать именно последнее значение („хорошая речь“), поскольку нас интересует качество речевого акта. (Дальнейшие рассуждения следуют из опыта Лаборатории фонетики Института языка и литературы Академии наук Латвийской ССР, а также из моих наблюдений в процессе преподавания сценической и ораторской речи.)

Качество речевого акта в основном определяется двумя факторами: 1) полнотой выражения содержания, 2. соответствием речевой ситуации. Содержание в свободной речи выражается полностью, т. е. выражается не только значение (в словах), но и смысл (о котором свидетельствует главным образом интонация). Следует учесть: если качество речи умень-

шается, прежде всего „пострадают“ средства для выражения смысла. С ситуацией свободная речь связана целесообразно, т. е. говорящий выполняет какую-то определенную задачу (информировать, доказать, убеждать и т. п.) и устраняет факторы, мешающие им (боязнь, плохое настроение и т. п.).

Таким образом — свободной назовем такую речь, которая 1. дает полную информацию о содержании и 2. целесообразно соответствует ситуации.

Разумеется, наша речь не всегда соответствует упомянутым требованиям. Довольно характерными в этом отношении являются следующие случаи. 1. Уменьшается выражение воли: говорящий не старается с помощью речи воздействовать на психику слушателя. 2. В интонации не выражается степень значимости и связь понятий (представлений) и вместе с тем языковых единиц. 3. В речи длительно выражается какая-то одна эмоция (грусть, радость и т. п.). В большинстве случаев имеет место не настоящее переживание, а плохая манера речи.

В речевой ситуации имеются некоторые элементы, в известной степени определяющие вероятность появления свободной речи, а именно: 1. импровизированный или репродуцированный текст, 2. обстановка повседневной или публичной речи, 3. степень общей подготовки говорящего (в частности его способность найти и выразить интерпретацию для данного текста).

Фонетист имеет дело чаще всего с 5 комбинациями этих элементов. Я перечисляю их в убывающем порядке вероятности свободной речи. 1. повседневная речь, 2. репродуцированная публичная речь хорошо подготовленных говорящих (напр., актеров), 3. импровизированная публичная речь хорошо подготовленных говорящих (напр., педагогов), 4. импровизированная публичная речь мало подготовленных говорящих, 5. репродуцированная публичная речь мало подготовленных говорящих.

Качество речевого акта должно быть особенно тщательно проанализировано, если он относится к пятой группе вероятности. Следует отметить, что именно из этой группы довольно часто берутся факты для фонетических исследований. Очевидно, более целесообразно использовать речевые акты первых трех групп.

DISCUSSION

Артемюв:

В лабораторных условиях возможно получить от диктора вполне естественную свободную речь, если, следуя методу К. С. Станиславского, выяснить с диктором: кто говорит, кому, что, с какой целью, в какой обстановке и с каким отношением.

К ВОПРОСУ ОБ АРТИКУЛЯЦИОННЫХ ОСОБЕННОСТЯХ АФФРИКАТ АРМЯНСКОГО ЯЗЫКА

А. А. ХАЧАТРЯН*

Консонантная система армянского языка характеризуется наличием трихотомических рядов смычных согласных, различающихся между собой наличием аспирации или голоса по отношению к глухому немаркированному члену этих оппозиций. В эту же систему включаются и аффрикаты, в точности повторяющие структуру смычных взрывных согласных.

В традиционной арменистике аффрикаты самостоятельно не классифицируются, а считаются сложными звуками, состоящими из взрывного (*t*), или аспирированного (*t'*), + щелинный ффрикативный (*s*) или (*š*) (или же их звонкие варианты).

Фонологический же анализ этих аффрикат показывает, что в системе армянского языка они функционируют как самостоятельные единицы, четко противопоставляемые в пределах слова соответствующим сочетаниям взрывной + ффрикативный, как, например, *kərtser* „младший“ и *kərcer* „вымя“ *patšar* „каменщик“ *patšar* „причина“. Гораздо больше случаев непротивопоставляемых сочетаний, в которых компоненты четко воспринимаются носителями языка как самостоятельные фонемы. Примеры: „достойный“ *katsa* „кастрюля“ *patžel* „наказать“. Исторически образованные от смычных согласных аффрикаты уже давно функционируют как самостоятельные фонемы, о чем свидетельствуют армянские письмена, созданные Месропом Маштоцом в IVв, и в которых эти фонемы представлены самостоятельными графемами. Поэтому функциональная самостоятельность аффрикат армянского языка не вызывает сомнений. Что же касается артикуляционно-акустической характеристики этих аффрикат, то здесь возникают серьезные разногласия, нуждающиеся в экспериментальной проверке. Используемая нами методика прямого палатографирования, рентгенографирования и кинорентгенографирования дает нам основание считать аффрикаты армянского языка особыми звуками, значительно отличными от соответствующих смычных и ффрикативных. Многочисленные палатограммы аффрикат, полученные от 20 дикторов и сопоставленные с палато-

* Институт языкознания АН Армянской ССР Ереван.

граммами взрывных (*t*), (*t'*), (*d*), показывают существенные различия. В то время, как при взрывных основная артикуляционная зона распространяется до передних резцов, образуя с внутренней поверхностью зубов довольно плотную смычку, при аффрикатах первой группы (*c*, *c'*, *j*) эта зона значительно сдвинута назад, охватывая лишь верхний край зубов и переднюю часть альвеол. При аффрикатах второй группы (*č*, *č'*, *ǰ*) эта зона еще дальше сдвинута назад, оставляя зубы совершенно незадетыми; площадь касания значительно шире, чем при (*t*), охватывая передний край твердого неба.

Характерной особенностью аффрикат как первой, так и второй группы является специфическая природа смычной фазы при их артикуляции, отличающейся неоднородностью, неустойчивостью качества смычки в отличие от чистых взрывных или фрикативных, при которых смычка резко очерчена и отличается большой контрастностью. При аффрикатах уже в стадии задержки происходит предвосхищение дальнейшей работы артикуляционных органов, а именно подготовка небольшой щели для протока воздуха. Положение этой щели на палатограммах намечается в виде ослабления окраски места смычки и сужением всей полосы смычки. Она внутри каждой группы аффрикат также неоднородна. Наиболее сильна она при глухих аффрикатах, слабее — у звонких, еще слабее — при аспирированных.

Сравнение рентгенограммы и лингвограммы аффрикат первой и второй групп подтверждает различия в артикуляционных зонах и работе языка. Активным органом при артикуляции первой группы аффрикат выступает кончик языка, при этом вся масса языка опущена, особенно в ее средней части, образуя значительную впадину. Ротовой резонатор при этом большой, а глоточный — сравнительно узкий. Артикуляция второй группы аффрикат характеризуется тем, что вся масса языка сдвинута вперед и значительно поднята вверх, создавая узкий ротовой резонатор, переходящий в широкий глоточный. Различия в артикуляции аффрикат еще более усиливаются включением губ в артикуляцию второй их группы. В то время, как при (*c*, *c'*, *j*) губы пассивно раскрыты, подготавливая артикуляцию последующей гласной, при (*č*, *č'*, *ǰ*) губы активно артикулируют, создавая широкий губной резонатор, что особенно четко видно из данных рентгена и кинорентгена.

Из всего сказанного вытекает, что уже на артикуляционном уровне аффрикатам армянского языка присущи специфические признаки, отличающие от соответствующих сочетаний взрывной + фрикативный. Этот вывод полностью согласуется с данными слухового анализа аффрикат при обратном прослушивании, проведенного нами со студентами и преподавателями армянского языка. Были проведены 3 серии опытов. Аудиторам в первой серии предъявлялись в магнитофонной записи инвертированные слоги типа ГС, где гласный оставался неизменным, а согласные (аффрикаты)

варьировались. Во второй серии опытов прослушивались слоги типа ГСГ, где гласные не изменялись, а аффрикаты варьировались. В третьей же серии были записаны сочетания типа фрикативный + взрывной в окружении гласных и слова, начитанные в обратном порядке, где аффрикаты как бы разложены на свои составляющие так, что при обратном воспроизведении они должны были прозвучать как нормальные слова. В опытах участвовало 64 аудитора.

Результаты первых двух серий опытов показали, что в аффрикатах фрикативный элемент опознавался правильно в 48 случаях, достаточно правильно (как оглушенный вариант) в 14 случаях, что составляло в целом 97 % распознавания. Смычный же элемент был распознан лишь в 16 случаях, что составляет 25 %. Скомпонованные же аффрикаты при инверсальном прослушивании звучали неестественно, как сочетания согласных, а не как единые звуки.

DISCUSSION

Артемюв:

Исследование Хачатрян продолжает славные традиции экспериментально-фонетического изучения звукового состава армянского языка, начатые еще Ачаряном в лаборатории Руссло.

Гаприндашвили:

Докладчица подробно остановилась на артикуляционные особенности аффрикат армянского языка, которые она изучила рентгенографическим и фотопалатографическим методами. Статистическая рентгенография, к сожалению, не может вскрыть всей полноты особенности динамики артикуляции исследуемых согласных, но данные, представленные в докладе очень ценны для изучения фонетических особенностей артикуляции аффрикат в армянском языке.

LE PHONÈME $[\tilde{\epsilon}]$ EN FRANÇAIS CANADIEN

RENÉ CHARBONNEAU*

Les travaux détaillés sur le français parlé au Canada, dont ceux de Pernot, Haden, Garner, et Locke, pour ne mentionner que ces noms, ne sont pas dépourvus d'intérêt.¹ Cependant les analyses qu'ils ont pratiquées sont à peu près toutes fondées sur des observations uniquement auditives. Il est peu de chercheurs, à l'exception de Haden, et Gendron qui aient eu recours aux méthodes d'analyse instrumentale.

Depuis quelques années, la radiocinématographie offre des possibilités des plus intéressantes et des plus efficaces pour observer et préciser une articulation. Lorsqu'on a réussi à synchroniser des films qui montrent les mouvements de la langue avec des films qui montrent les mouvements des lèvres et qu'on peut entendre simultanément le son correspondant (avec sonagrammes), il ne fait pas de doute qu'on est en présence de documents de première valeur pour décrire les différents aspects physiologique et acoustique d'une voyelle. Un article de M. P. Delattre intitulé: „Voyelles pures et voyelles diphtonguées“, publié dans *French Review* en 1963 et repris dans une publication plus récente chez Mouton,² est à ce point de vue très révélateur et riche de détails.

On reste toujours frappé par les caractères bien particuliers qui se dégagent de l'audition des voyelles nasales prononcées par un Canadien français. Sans doute les quatre phonèmes $[\tilde{a}/\tilde{\epsilon}/\tilde{o}/\tilde{\omega}]$ du français général existent-ils au Canada, mais le timbre de ces voyelles est différent de celui qu'on distingue habituellement en France dans la prononciation normale.

Dans le livre qu'il vient de publier chez Klincksieck et aux Presses de l'Université Laval, à Québec, notre collègue et ami, M. Jean-Denis Gendron, écrit ce qui suit au chapitre qui traite des voyelles nasales: „Il est évident que l'articulation des voyelles nasales et, par conséquent, leur aspect auditif ne sont pas les mêmes dans le français parlé au Canada et dans celui de Paris, voire dans la langue générale. La différence est surtout frappante si l'on écoute le parler canadien populaire, mais elle existe aussi dans la prononciation des milieux instruits. Ceux-ci tendent à se rapprocher

* Département de linguistique et de langues modernes, Université de Montréal

¹ Gendron, Jean-Denis, *Tendances phonétiques du français parlé au Canada*, Klincksieck, Paris — Les Presses de l'Université Laval, Québec, 1966: Voir introduction.

² Delattre, Pierre, *Studies in French and Comparative Phonetics*, Mouton, 1966: p. 95.

de la norme, mais... on n'y réussit pas toujours... et le [ɛ̃] semble être de toutes les voyelles nasales le moins modifié dans le sens de la correction orthophonique³ Voici un échantillon de prononciation populaire.

[Ruban magnétique...]

Notre intention est de vous entretenir brièvement de cette voyelle nasale dont l'étude est basée sur la radiocinématographie et la spectrographie. Un premier travail entrepris à l'Université de Strasbourg sous la direction de M. Straka et avec l'excellente collaboration de Madame Simon nous permet de vous faire connaître les premiers résultats entrevus.

Nous avons fait exécuter deux films de 180 mètres chacun (deux sujets) et nous avons totalisé un ensemble de 648 phrases comportant des voyelles nasales en différentes positions. Chaque phrase était formée de quatre syllabes. Les films ont été tournés à la vitesse de 36 images à la seconde, chaque image représentant ainsi une durée de 2.8 cs. Le champ d'observation de l'appareil, de 23 centimètres, permettait de circonscrire l'espace compris entre les lèvres et la paroi pharyngale sur le plan horizontal, entre l'os hyoïde et le dessus de l'os palatal sur le plan vertical. Comme notre étude se limitait plus spécialement à observer l'activité de la langue et du voile du palais, nous avons dû, pour ne rien laisser échapper, sacrifier parfois la visibilité des lèvres, la luminosité se concentrant davantage sur les cavités buccale et pharyngale. L'activité des lèvres reste cependant visible et nous procédons depuis quelque temps à en préciser les détails. Nous avons aussi établi des comparaisons avec la voyelle orale correspondante. Les dimensions des cavités de résonance ont été établies à partir des 7 points suivants: angle des maxillaires, langue région-alvéolaire, palatale, sommet de la voûte; langue-voile; langue-paroi pharyngale; voile-paroi pharyngale.

Nous avons aussi procédé à l'analyse acoustique de ces voyelles à l'aide du Sonagraph, en prenant pour acquis que les deux facteurs de la nasalité sont: la réduction de l'intensité de ce qui est le 1^{er} formant dans les voyelles orales (facteur de beaucoup le plus important) et un formant additionnel étroit et faible à environ 250 cycles. Nous avons considéré les autres facteurs comme inutiles, car ils ne produisent aucun effet de nasalisation à l'oreille. En dessus de 2500 cycles, les formants ne sont plus linguistiques, puisqu'ils ne changent plus les phonèmes, mais seulement la qualité individuelle de la voix. Les sonagrammes, comme nous le verrons, fournissent des renseignements utiles sur la stabilité auditive de la voyelle. Ils exigent cependant une grande attention dans la lecture qui permettra d'éviter la confusion entre le changement continu du timbre d'une voyelle et les transitions des formants qui sont en relation étroite avec la perception de la consonne. Voici quelques exemples de [ɛ̃] en syllabe ouverte et en syllabe fermée accentuée:

³ Gendron, Jean—Denis, *Tendances phonétiques du français parlé au Canada*, Klincksieck, Paris — Les Presses de l'Université Laval, Québec, 1966: p. 106.

[ɛ̃] DANS LE MOT PAIN:

Après l'abandon par les lèvres (pas dessinées ici) de la tenue de l'occlusive, la langue recule progressivement et, en s'abaissant dans sa partie antérieure, elle s'élève en même temps dans sa partie postérieure (phases 1 à 5). Simultanément, l'angle des maxillaires s'agrandit plus particulièrement des phases 3 à 5. Ici semble atteint le segment central de la voyelle. C'est à ce stade que l'angle des maxillaires et la largeur du canal buccal atteignent leur maximum. Les cavités ne cesseront ensuite de diminuer en arrière comme en avant. Le [ɛ̃] est relativement ouvert au départ dans la région palatale. Il est très postérieur (beaucoup plus qu'en français) et le dos de la langue remonte sensiblement en direction du voile, produisant à cet endroit un passage étroit inaccoutumé dans la formation des voyelles françaises correspondantes. Par la suite, la largeur de cette cavité se modifiera en fonction du degré d'élévation de la langue et de ses déplacements sur le plan horizontal. — Le voile esquisse, à la phase no 3, un léger éloignement de la paroi rhino-pharyngale, mais le détachement est pratiqué plutôt brusquement à la phase no 5, et le passage qu'il produit indique l'absence d'une pleine nasalité. Il faudra attendre quelques centisecondes pour l'obtention d'une résonance nasale complète. La langue continue sa trajectoire en reculant légèrement et en s'élevant considérablement dans sa partie centrale en direction du sommet de la voûte. Ce mouvement a ses conséquences: élargissement du résonateur pharyngal et diminution de l'espace compris entre la langue et la région alvéolaire, la langue et la région palatale. Contrairement à ces modifications d'aperture, l'angle des maxillaires est relativement stable des phases 5 à 10.

Puis un changement essentiel apparaît en arrière. Tout en maintenant sa position haute, la masse de la langue rencontre la partie centrale du voile. Mais il y a hésitation, flottement. La solidité du contact n'est pas assurée. Ce n'est qu'à l'image suivante qu'il y a attouchement.

En résumé, pour le [ɛ̃] du français canadien en syllabe libre accentuée, les organes articulatoires accusent des mouvements tout à fait particuliers: au départ, c'est une voyelle relativement ouverte dans le résonateur avant, fermée dans le résonateur arrière; puis avec le mouvement d'élévation de la langue, c'est le contraire qui se produit: rétrécissement du résonateur avant et élargissement du résonateur arrière. Enfin, rapprochement du dos de la langue et du voile.

Le [ɛ̃] est nasillard. Il donne l'impression auditive d'être prononcé très en avant. Pourtant, lorsqu'il se dénasalise, on entend un [ɛ] plutôt qu'un [e]. „Un verre de vin“, pour une oreille non habituée, est entendu quelquefois „un verre de lait“, cette dénasalisation portant à confusion. Par ailleurs, si on tient compte de l'élévation de la langue, de la largeur des dimensions des résonateurs buccal et pharyngal, c'est plutôt un [e] englobé d'une résonance nasale qui se dirige et se perd vers l'aigu.

AUTRES EXEMPLES:

Planches radio	003	c'est un requin
	004	vous avez faim
	005	à la Toussaint
Sonagrammes	x	fiche-moi la paix
	y	le temps passait
	z	c'est bien mal fait
	001	donne-lui du pain
	003	c'est un requin
	005	à la Toussaint

Si on considère le [ɛ̃] — en position de syllabe fermée accentuée —, la durée joue un rôle important dans la modification du timbre de la voyelle.

Soit [ɛ̃] dans [fɛ̃:t].

Depuis la phase no 5, où l'angle des maxillaires atteint un maximum, la partie antérieure de la langue qui avance progressivement en direction de la région palatale, décrit un mouvement circulaire, passe par une gamme de positions successives qui font varier le timbre. L'abandon du contact — voile — paroi rhino-pharyngale — se produit au stade no 3. C'est la diphtongaison de [ɛ̃] en position accentuée antéconsonantique déjà pressentie par les phonéticiens cités précédemment et qui demeure une des préoccupations des orthophonistes. Elle s'entend surtout dans la prononciation populaire, rarement ou de façon très atténuée dans celle des classes cultivées: [fɛ̃:t], [kɛ̃:t], [pɛ̃:t], [sɛ̃:t].

Autres planches Radio	007	teinte
	006	pinte
Sonagrammes	260	fête
	163	tête
	009	feinte
	007	teinte

Les voyelles nasales très fermées du français canadien sont-elles phonologiquement différentes des voyelles nasales françaises? Certes, elles se distinguent par leur timbre plus aigu, leur nasalité moindre (au début de l'émission) et leur durée plus longue, mais ces différences ne perturbent pas les rapports entre ces voyelles et, par conséquent, le système phonologique de la langue. Elles représentent des réalisations articulatoires qui ne modifient pas les oppositions phonémiques établies et ne créent pas à l'intérieur du français canadien des confusions d'ordre sémantique. — Il en est parfois autrement par rapport au français de France, et un Français écoutant un Canadien français ne sera pas seulement étonné par des différences de prononciation

qu'il qualifiera vaguement d'accent, de nasillement, de "méridionalisme", mais il pourra confondre occasionnellement, comme nous l'avons déjà dit, une voyelle nasale franco-canadienne avec une voyelle orale française. De plus, la diphtongaison des voyelles nasales du français canadien ne compromet pas les rapports entre les systèmes des voyelles nasales des deux parlers. Certes, on pourrait se demander si la voyelle nasale canadienne est, par rapport à la voyelle nasale française correspondante, une réalité à traits distinctifs multiples, mais la commutation nous aide à résoudre cette question: les différences entre la voyelle canadienne et la voyelle française sont plutôt d'ordre allophonique; elles forment un tout indissociable.

Le [ɛ̃] entendu au Canada français dans la classe populaire demeure-t-il le vestige d'un archaïsme, d'une prononciation ancienne transmise et conservée jusqu'à nos jours? Il serait intéressant de faire des recherches sur ce sujet et d'établir si, dans certaines régions de Bretagne, de Normandie, etc., lieux d'origine de nos premiers Canadiens, se retrouve cette particularité articulatoire que nous venons d'esquisser.

DISCUSSION

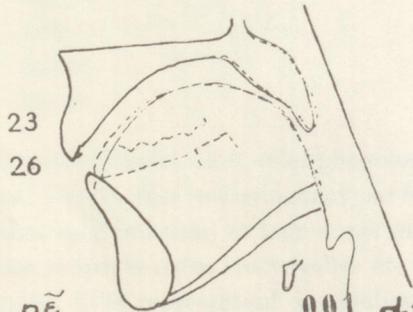
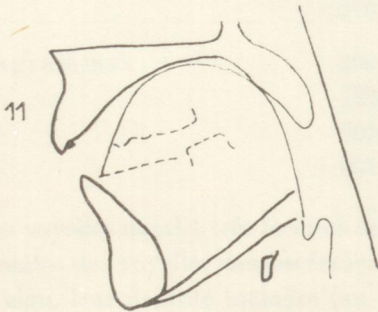
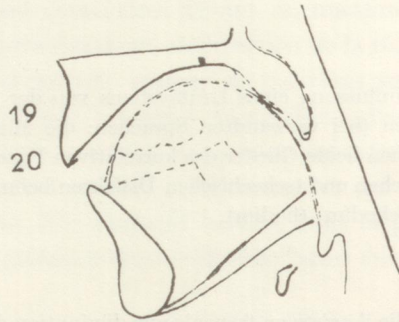
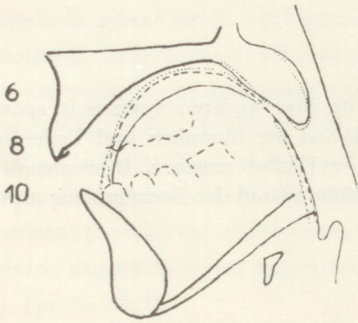
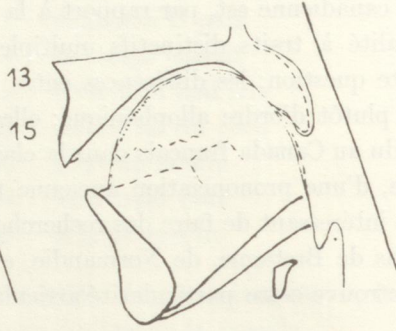
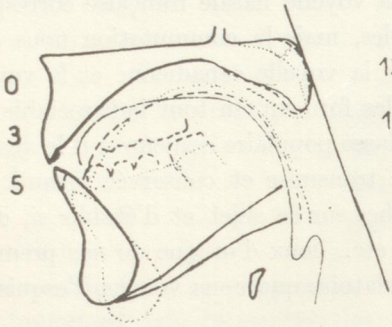
Blanár:

Bei der Beeinflussung eines Lautsystems von der Seite eines anderen Systems in speziellen Sprachkontakten (bei verwandten Sprachen: die Mundarten der Slowaken und Tschechen in Bulgarien) können beide Glieder des korrelativen Paares den Einfluß zeigen. Z. B. werden (h—ch) in den slowakischen und tschechischen Dialekten beim größten Grad der Beeinflussung nicht gut unterschieden (chodím) (hodím).

Haudricourt:

En Normandie il existe en français une diminution de longueur pour les voyelles nasales: tente [tāt] ~ tante [tā:t] voir Word (1954) vérifié depuis par des témoignages vivants.

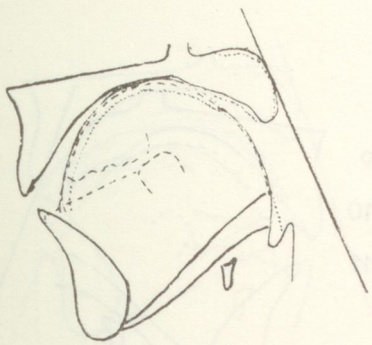
PLANCHE - RADIO



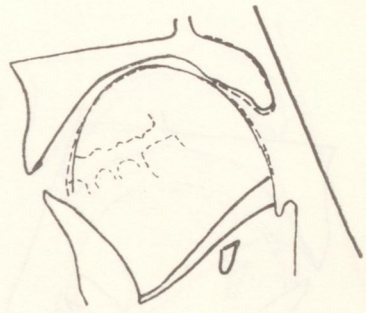
001 don lyi dy pē

001 don lyi
PĒ

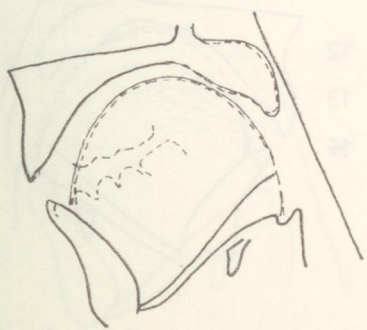
0
1
2



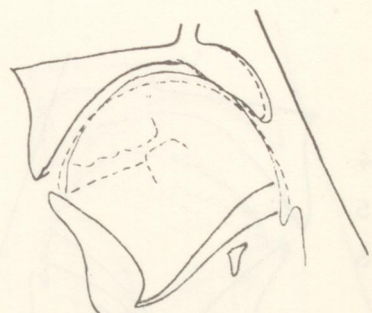
8
10



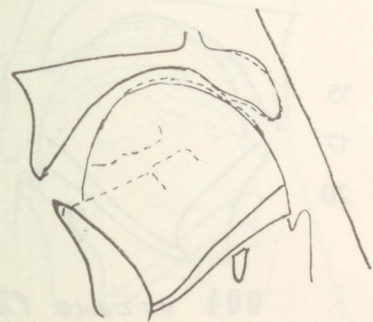
3
4



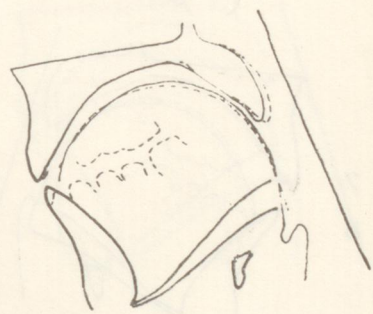
12
17



5
7



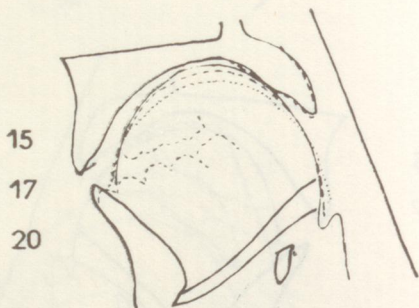
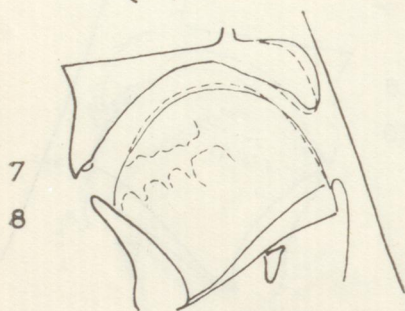
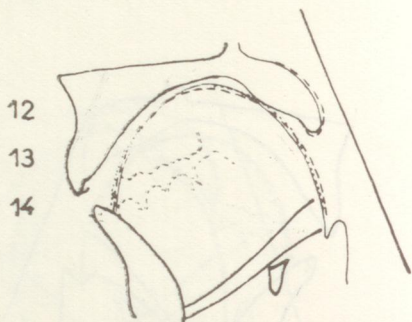
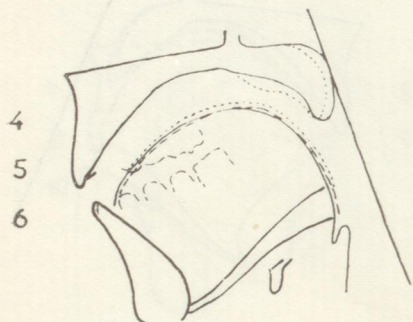
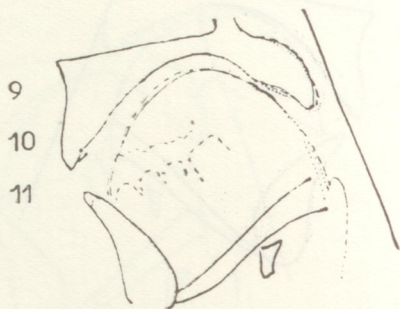
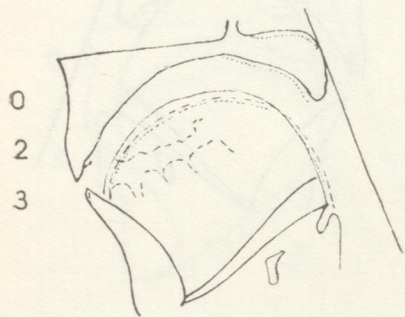
18
20



002 se tõe RAKĚ

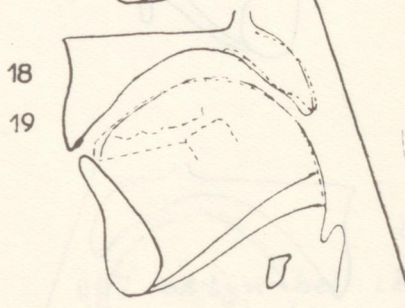
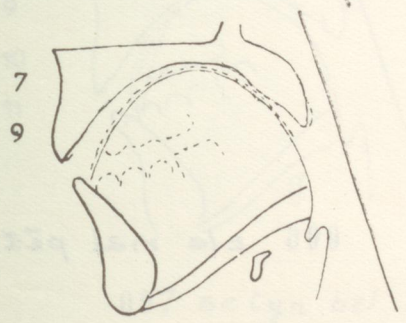
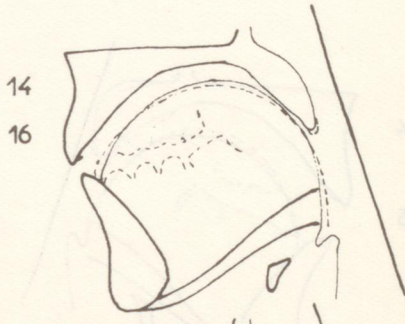
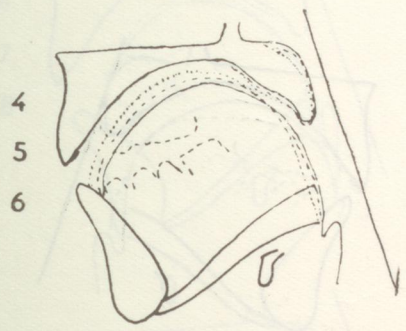
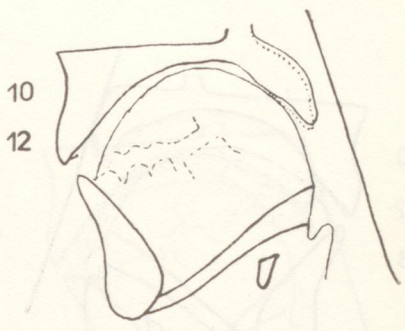
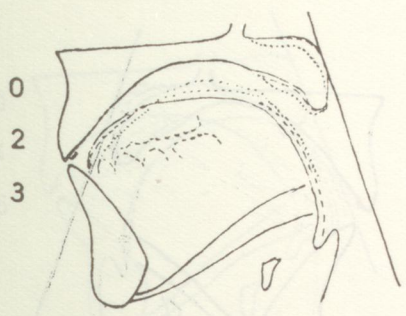
003 se tõe RAKĚ

Fig. 1. Planche Radio.



004 vuzave fĚ

004 vuzave fĚ

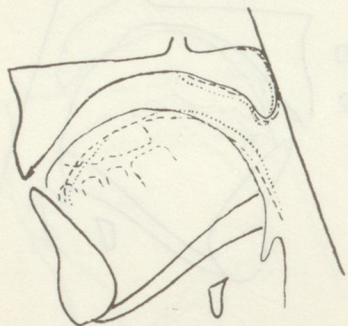


005 a la tuse

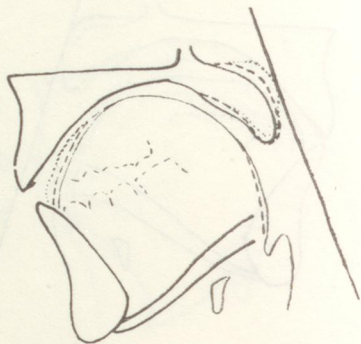
005 a la tuse

Fig 2.

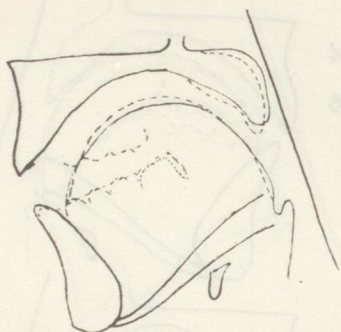
0
2
3



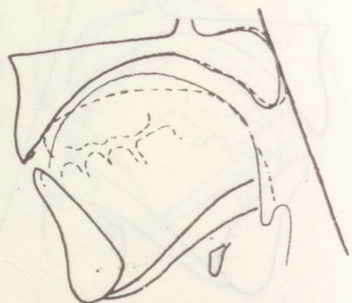
9
10
11



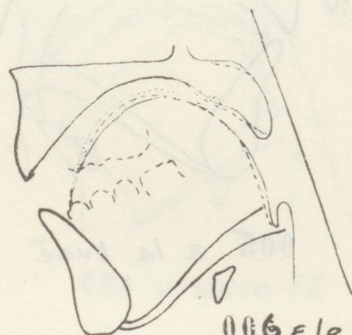
4
5



12
14

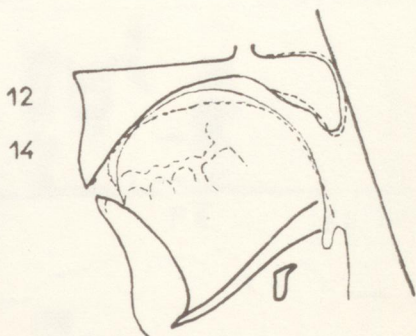
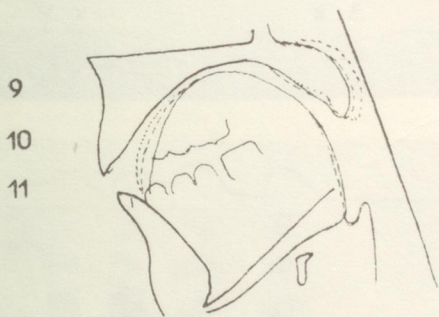
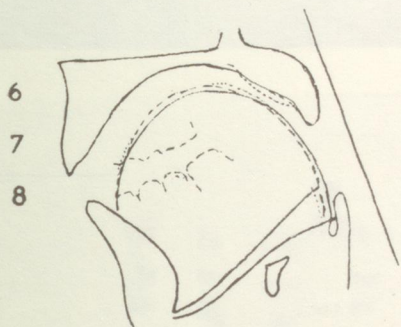
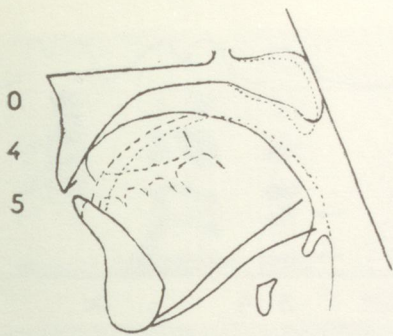


6
7
8



006 e/e mal pē:t

006 e/e mal pē:t



007 setyn. bel tēt

007 setyn bel tēt

Fig. 3.

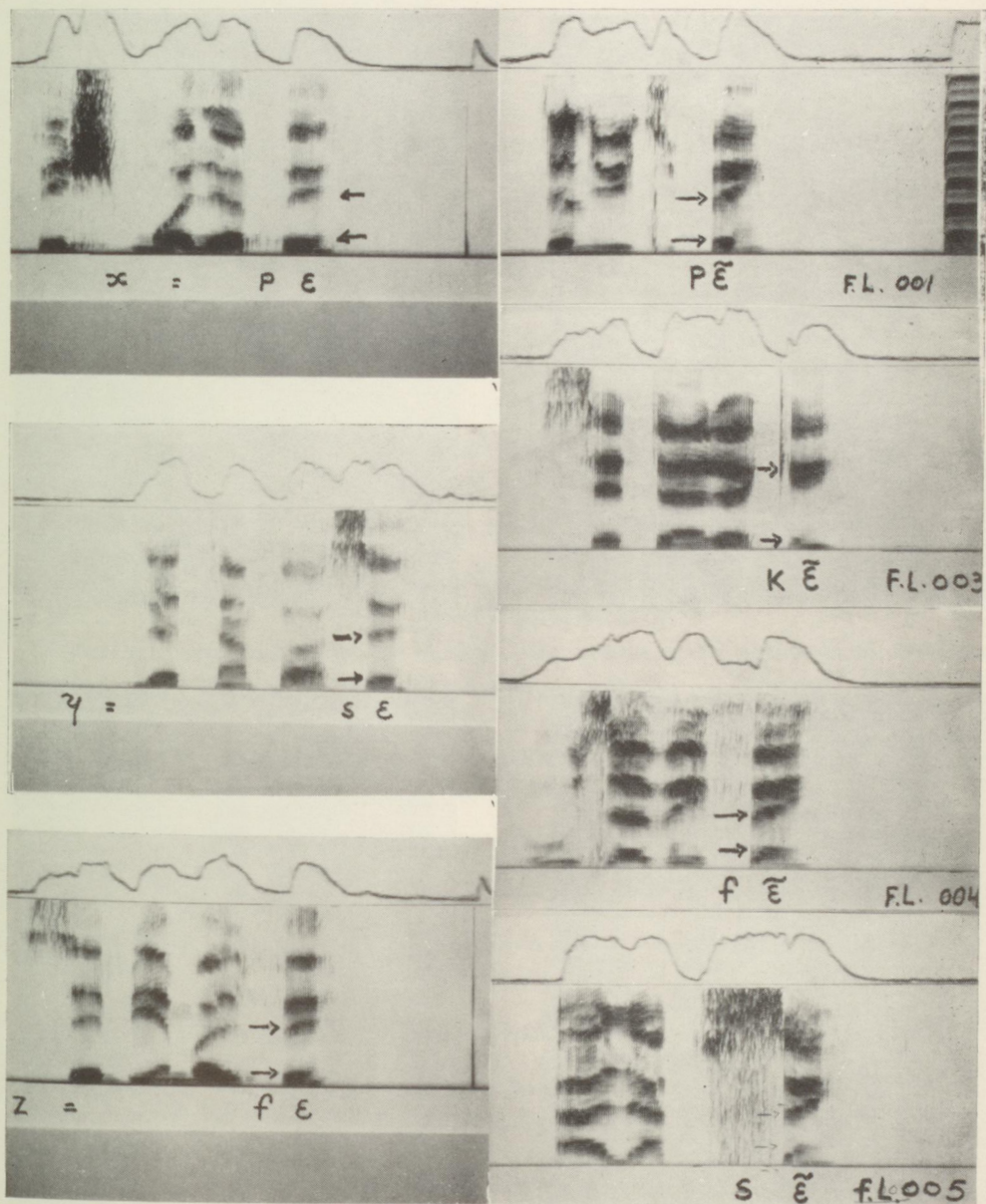


Fig. 4.

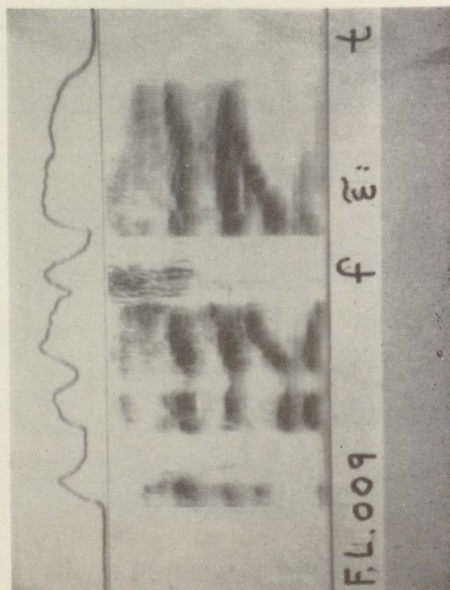
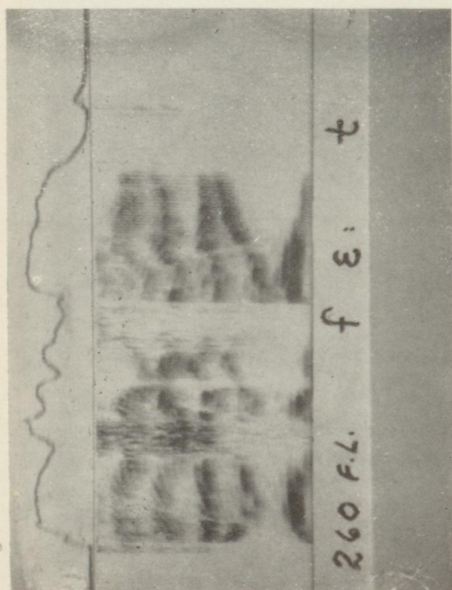
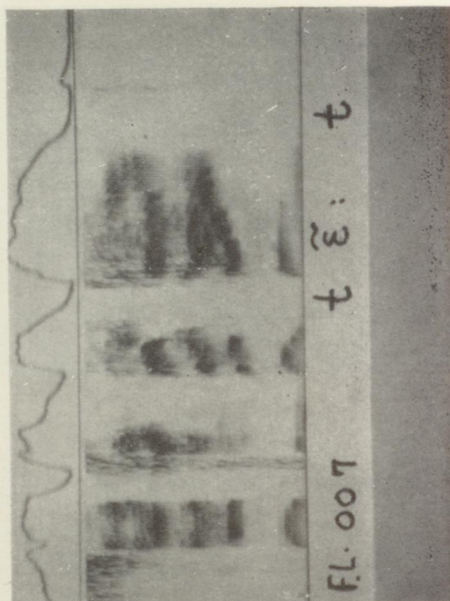
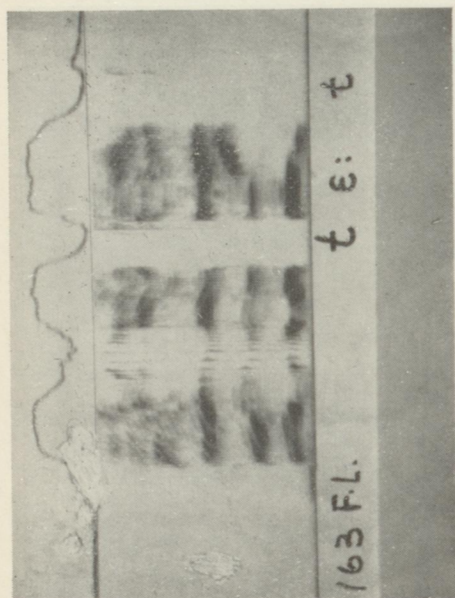


Fig. 5.

A STATISTICAL APPROACH TO SPEECH ANALYSIS BY WAVE FORM

PHILIPP CHRISTOV*

Let us take as model for the speech source a linear dynamic system characterized by a memory function which, driven by a forcing function, gives the speech signal at its output.¹ The forcing function is a quasiperiodic train of laryngeal excitations which follow one after another, by $t > 0$, in intervals equal to the pitch period² and have so short a duration that at the end of each pitch period equal zero.³ As the memory function seems to have such characteristics that at the end of each pitch period the response of each single glottal pulse also equals zero,⁴ the latter can be expressed with the 'identical' functions⁵ or *elementary waves*

$$x_k(\tau; a_k) = \begin{cases} x_k(\tau; a_k) & \text{if } t_k - T_{ok} \leq \tau < t_k \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where \mathbf{a} is an aggregate of waveform parameters, T_{ok} is the value of the pitch 'period' in the k^{th} elementary wave and k is chosen arbitrarily.

From the speculation about the nature of the speech signal, made above, it follows that each single elementary wave contains no periodic components and is statistically independent. Hence each vocal segment of the speech signal, defined as a *phone*⁶ (Fig. 1), can be regarded as a non-stationary *random function* $X(\tau; \mathbf{a})$ ⁷ or *random speech wave* presented by the set of its trials

$$x_1(\tau; a_1), x_2(\tau; a_2), x_3(\tau; a_3) \dots x_k(\tau; a_k), \dots x_n(\tau; a_n) \quad (2)$$

* Bulgarian Language Institute, Sofia, Bulgaria.

¹ Huggins, W. H., "A Note in Autocorrelation Analysis of Speech Sounds", *JASA*, **26**, 790—794 (1954).

² Fant, G., "Acoustic Theory of Speech Production", перевод с английского под редакцией В. С. Григорьева, из-во Наука, Москва (1964), стр. 29.

³ Miller, R. L., "Nature of the Vocal Cord Wave", *JASA*, **31**, 667—677 (1959).

⁴ Drew, R. O., and Kellog, E. W., "Starting Characteristics of Speech Sounds", *JASA*, **12**, 95—103 (1940), p. 103.

⁵ Поливанов, К. М., "Теоретические основы электротехники", част I, из-во Энергия, Москва, Ленинград (1965), стр. 374.

⁶ Bloch, B., "Studies in Colloquial Japanese", *Language*, **26**, 86 (1950), p. 89.

⁷ Middleton, D., "An Introduction to the Statistical Communication Theory", перевод с английского, том I, из-во Советское радио, Москва (1961), стр. 676.

which occur in the fixed interval $(\tau_0, \tau_0 + T_0)$, where $T_0 = C^{te} \geq T_{0k}$ and whereas k is chosen arbitrarily, it may be chosen in an orderly manner as well.

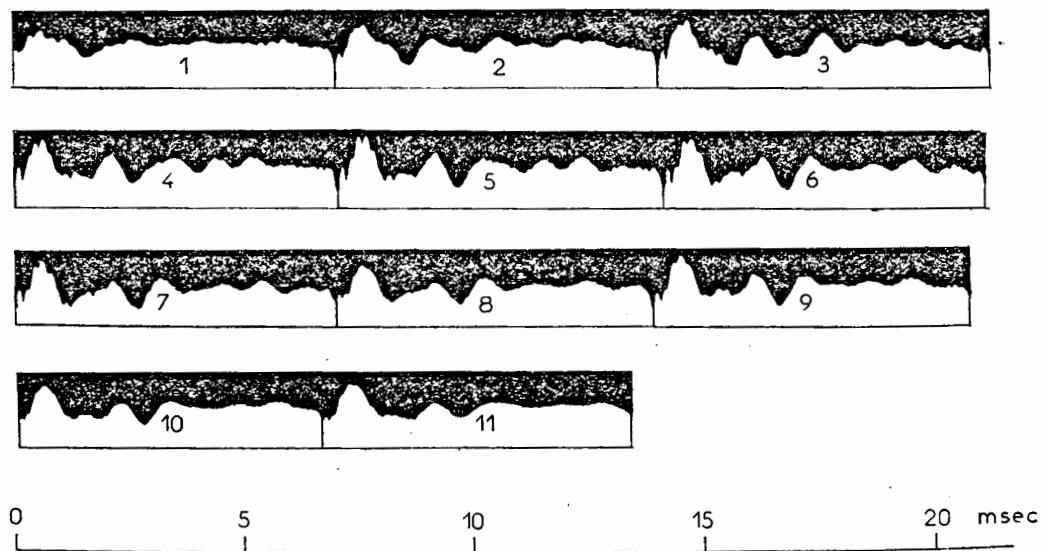


Fig. 1. High speed oscillogram of the vowel [á] as in [dáp]. Speaker A. M. The oscillogram is obtained by photographic enlargement of bilateral motion picture sound recording.

Let us take samples $x_k(\tau_m; a_k) \equiv x_{km}$ of the k^{th} elementary wave in its successive phases of development $\tau_m [m = 0, 1, 2, 3, \dots n (k)]$ at constant sampling intervals $\Delta\tau = \tau_m - \tau_{m-1} = 1/2W_x$, where W_x is the signal band width, which are sufficiently long to ensure the independence of the samples.⁸ Then each elementary wave is presented with the set of samples

$$x_k^*(\tau; a_k) \equiv x_{k0}, x_{k1}, x_{k2}, x_{k3}, \dots, x_{km}, \dots, x_{kn(k)}, \dots, x_{kn} \quad (3)$$

where the samples x_{k0} in the points of zero-crossing and the samples with numbers of sampling between $n(k) = T_{0k}/\Delta\tau$ and $n = T_0/\Delta\tau$ (See Eq. 1) equals zero.

If the sections of the random speech wave $X(\tau; \mathbf{a})$ in the fixed phases of development τ_m are grouped to form ensembles

$$X(\tau_m; \mathbf{a}) = x_{1m}, x_{2m}, x_{3m}, \dots, x_{km}, \dots, x_{hm} \quad (4)$$

then the ensembles or random variables $X(\tau_m; \mathbf{a})$ form the system of random variables $X^*(\tau; \mathbf{a})$

$$X^*(\tau; \mathbf{a}) \equiv [X(\tau_1; \mathbf{a}), X(\tau_2; \mathbf{a}), X(\tau_3; \mathbf{a}), \dots, X(\tau_m; \mathbf{a}), \dots, X(\tau_n; \mathbf{a})] \quad (5)$$

⁸ Lee, Y. W., "Statistical Theory of Communication", Fourth printing, John Wiley & Sons, Inc., New York, London, Sydney (1946), p. 278.

Let us assume that the storage rule for the information about the system of random variables $X^*(\tau; \mathbf{a})$ is given by the rectangular matrix χ

$$\chi = \|x_{km}\| (k = 1, 2, 3, 4, \dots, h) (m = 1, 2, 3, 4, \dots, n(k), \dots, n) \quad (6)$$

Then we can consider as a portrayal of the system of random variables the discrete vector field created in the measurement space as a result of the storage operation, i.e., as a discrete vector variable $\mathbf{X}^* = f\chi(\mathbf{Q})$, where the operator, $f\chi$ designates that \mathbf{X}^* takes its values in the points \mathbf{Q} of the measurement space, according to the rules given by the matrix χ .

Assuming a measurement space determined by amplitude, phase and time, $\mathbf{Q}(x, \tau, t)$, we can decompose the vector variable χ^* in the components belonging to the subspaces of the measurement space (Fig. 2):

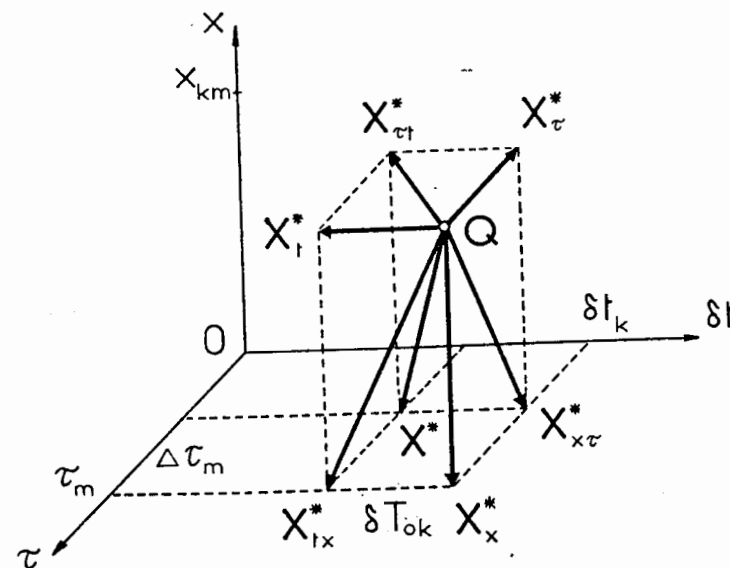


Fig. 2. The discrete vector variable $\mathbf{X}^* = f\chi(\mathbf{Q})$ and its components in the point (x_{km}, τ_m, t_k) of the measurement space, determined by amplitude x , phase τ and time t ($\delta = C^{te}$ is scale reduction coefficient).

1. *Unidimensional Subspaces.* $\mathbf{X}^* = \mathbf{X}_x^* + \mathbf{X}_\tau^* + \mathbf{X}_t^*$ where $\mathbf{X}_x^* = -x_{km}x_x^0$, $\mathbf{X}_\tau^* = -\Delta\tau x_\tau^0$ and $\mathbf{X}_t^* = -\delta T_{0k}x_t^0$ and where $\delta = C^{te}$ is scale reduction coefficient which takes values between zero and unity, $0 < \delta \leq 1$. The unidimensional components form sets of measurements which generates the statistical parameters (mean, variance, etc.) of the distributions of amplitude and pitch 'period', and the duration of the speech sound. The space filling properties of the portrayal \mathbf{X}^* are given by the mixed product $V_{\mathbf{X}^*} = \mathbf{X}_x^* \cdot (\mathbf{X}_\tau^* \times \mathbf{X}_t^*)$. The unit vector $\mathbf{x}^0 = \mathbf{X}^*/|\mathbf{X}^*|$ of \mathbf{X}^* gives its directional properties.

2. *Two Dimensional Subspaces.* $\mathbf{X}^* = \mathbf{X}_{x\tau}^* + \mathbf{X}_{\tau t}^* + \mathbf{X}_{t x}^*$, which generate a set of subportrayals: the plot $\bar{x}_p(t)$ of mean peak amplitude ($\bar{x}_p = x_{p-p/2}$) vs time or mean

amplitude envelope, the plot $T_{0(t)}$ of pitch 'period' vs time and the family of waveforms of the trials of the random function.

It is known,⁹ that if we let the total number of samplings per speech sound tend to discontinuity then the system of random variables (Eq. 5), presented by the pattern \mathbf{X}^* , becomes equivalent to the random speech wave $X(\tau, \mathbf{a})$. Hence any visual representation of speech based on this principle^{10,11} should be considered as reliable (Fk 3).

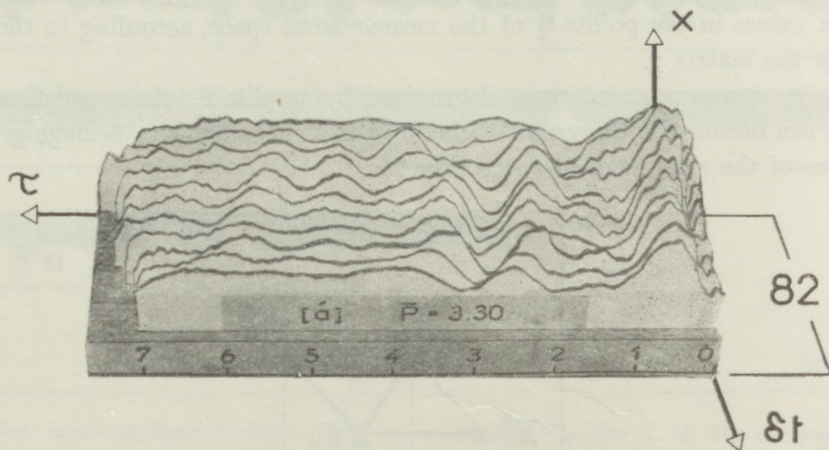


Fig. 3. Waveform (WF) Portrayal of the vowel [á] as in [dáp] build from the high speed oscillogram shown in Fig. 1. Total number of elementary waves $h = 11$. Recorder scale [sec. 10^{-3}]. Duration $82 \cdot 10^{-3}$ sec ($\delta = 0,032$).

The pitch 'period' normalized portrayal \mathbf{B}^* of the random function generates its mean, variance, correlation function and power spectra. Assuming that each trial $x_k(\tau; \mathbf{a}_k)$ exists between two successive positive-going zero-crossings, the portrayal \mathbf{B}^* can be obtained after prearranging of the time-domain samples x_{km} in a new phase-angle basis according to the rule given by the matrix \mathbf{B}

$$\mathbf{B} = \|\mathbf{B}_{km}\| \quad (k = 1, 2, 3 \dots h) \quad (M = 1, 2, 3 \dots N)$$

where $\mathbf{B}_{km} = \|x_{km} \dots x_{k(m+\vartheta_k-1)}\|$
 and where $\pm \vartheta_k \approx \pm n(k)/2\pi$ is an entire digit.

But we may choose another way for analysis of the waveform portrayal by which no pitch normalization is needed; we can use parameters for direct evaluation of the successive waveforms like the *crest factor* C and the *form factor* F . These parameters, together with the *slope of the pulse front* $S = \text{Peak/Rise time}$,

⁹) Пугачев, В. С., „Теория случайных функций“, Физматгиз, Москва (1960), p. 204.

¹⁰ Christov, P., „New Methods for Investigation of Speech Sounds“, *Technika*, **11** 215—218 (1962) (In Bulgarian).

¹¹ Grützmaier, M., „Demonstration eines Tonhörenschrifters“, *Phonetica*, **13**, 3—17 (1965).

can be applied also in regard to the mean envelope $\bar{x}_p(t)$ and to the plots of the another waveform parameters *vs* time: $T_0(t)$, $C(t)$, $F(t)$, etc.

Some preliminary results of the application of the methods described in this paper to the practical problems of speech analysis can be considered as encouraging. Five stressed Bulgarian vowels, uttered by speaker A. M., in nonsense syllables, was subjected to speech analysis by the waveform. Sampling with $\Delta\tau = 0,0001$ sec was

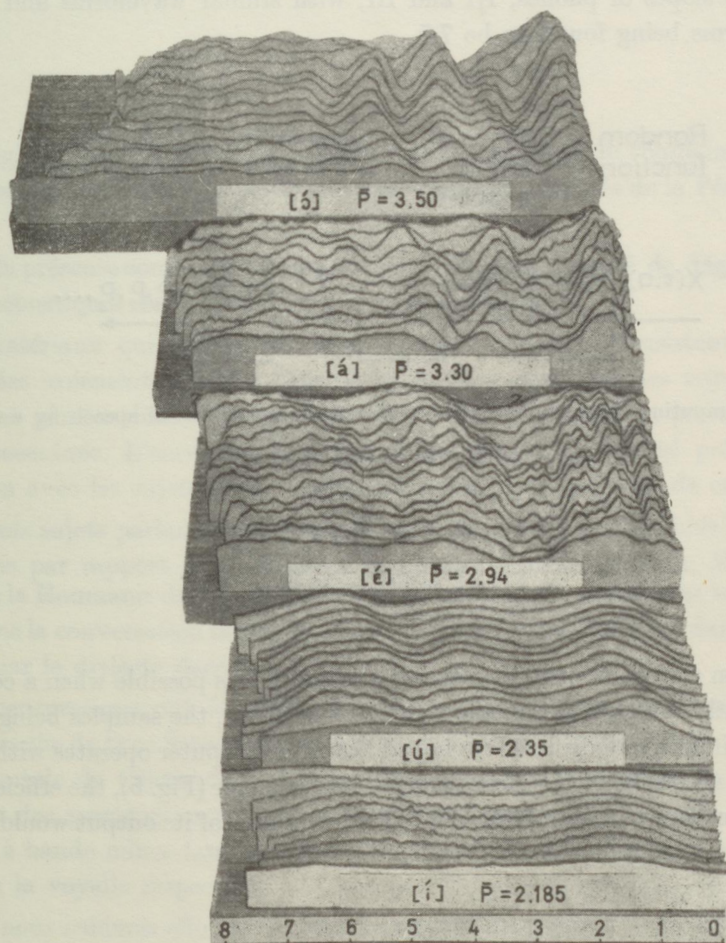


Fig. 4. WF Portrayals of five vowels uttered by speaker A. M., in the syllables: $[dik]$, $[pút]$, $[t/ét]$, $[dáp]$ and $[tón]$ (Phonetic designations according to I.P.A.). Recorder scale [sec. 10^{-3}].

accomplished manually from their Waveform (WF) Portrayals shown in Fig. 4. The mean product-factor of the waveform $\bar{P} = \bar{C} \cdot \bar{F}$, presented in the right-hand side of Fig. 4., seems to match the complexity of the waveform and, since it is non-dimensional, it appears that it is closely correlated to the phonemic value of these vowels. The observation that the sudden change in voice effort during the stress

results in corresponding changes in the waveform¹² makes it reasonable to suggest that the product $(C_{P(t)} \cdot C_{x_p(t)})$ would be effective by digital evaluation of stress. The envelope form of IrI¹³ can be evaluated by the product factor of the envelope $P_{x_p(t)}$, the ratio between its value in the modulated and non-modulated segments of the carrier vowel sound being greater or equal to 1,55. It is suggested that the most important acoustic cue of IjI is the slope $S_{T_0(t)}$ of the plot of pitch vs time,¹⁴ the ratio between the slopes of phones, IjI and IiI, with similar waveforms and amplitude envelope forms being found to be 7,5.

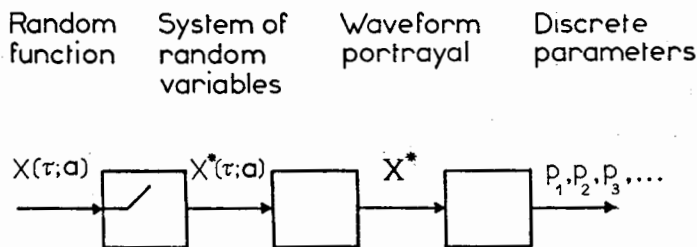


Fig. 5. Computing circuit of the process of statistical analysis of speech by waveform.

$$\begin{aligned}
 & Q(x_{km}, \tau_m, t_k) \\
 & \mathbf{X}^* = -x_{km} \mathbf{x}_x^0 \\
 & \mathbf{X}_\tau^* = -\frac{\tau m}{m} \mathbf{x}_\tau^0 \\
 & \mathbf{X}_t^* = \delta(t_{k-1} - t) \mathbf{x}_t^0
 \end{aligned}$$

It has been shown recently that speech recognition is possible when a computer is presented with short samples of the acoustic waveform, the samples being processed without preliminary analysis.¹⁵ It is hoped that if a computer operates with the input speech wave according to the ideas set forth in this paper (Fig. 5), the efficiency of the process of mechanical recognition as well as the quality of its output would be improved.

¹² Christov, P., "Some Peculiarities of the Bulgarian Vowels", *Bulgarian Language*, **13**, 22—27 (1963), p. 26 (In Bulgarian. French abstract in Bull., Anyl., Litt., Sci., Bulg., A. 1/1963, No 21).

¹³ Christov, P., "Experiment for Changing the Envelope Form of Vowels", 5^e C.I.A., *Reports*, **1**^a, Liège (1965), A11.

¹⁴ Christov, P., "Investigation Upon the Sonants in the Bulgarian Language", *Bulgarian Language*, **14**, 32—39 (1964), p. 37 (In Bulgarian. French abstract in Bull., Anal., Litt., Sci., Bulg., A. 1/1964, No 8).

¹⁵ Reddy, D. R., "Segmentation of Speech Sounds", *JASA*, **40**, 307—312 (1966).

LES VOYELLES MÉGLÉNO-ROUMAINES

CORNELIA COHUȚ—MARIA MĂRDĂRESCU

Le mégléno-roumain, l'une des quatre grandes divisions du roumain, est parlé par un groupe peu nombreux de personnes dans quelques localités de la Péninsule Balkanique.

Dans la présente communication nous nous sommes proposé de dégager les traits physicoacoustiques des voyelles du mégléno-roumain.

Les matériaux qui forment la base de notre recherche consistent des voyelles prononcées isolément et d'une liste de 70 mots comprenant des voyelles mégléno-roumaines placées dans certains contextes consonnantiques en position accentuée et non-accentuée. L'enregistrement sur ruban magnétique a été précédé par une discussion avec les sujets, pour vérifier l'exactitude de la forme de ces mots.

Les trois sujets parlants (des hommes) sont originaires des localités les plus représentatives par rapport aux différents patois du mégléno-roumain. Mais ces sujets habitant la Roumanie depuis 40 ans environ, leur parler — dont ils se servent uniquement dans la conversation familière — a été probablement influencé dans une certaine mesure par le dialecte daco-roumain.

Les mensurations ont été faites sur des spectrogrammes réalisés à l'aide d'un spectrographe de type Sonagraph Kay Electric Co., dans le laboratoire de phonétique expérimentale du Centre de recherches phonétiques et dialectales de Bucarest. On a obtenu des spectrogrammes proprement dits à bande large („wide band“) et des sections à bande mince („narrow band“) — généralement effectuées au milieu de la durée de la voyelle respective.

Nous nous sommes efforcés — dans la mesure du possible — de déterminer pour chaque voyelle les fréquences des trois premiers formants.

Nous avons comparé nos résultats aux données antérieurement obtenues dans l'étude acoustique des voyelles du roumain littéraire (c'est-à-dire de l'aspect littéraire du dialecte daco-roumain)¹ ainsi qu'aux voyelles existantes seulement dans certains

¹ Voir Andrei Avram, *Vocalele românești [ă] și [î] din punct de vedere acustic*, *Studii și cercetări lingvistice*, XIV, 1963, 2, p. 165—177; Valeriu Șuteu, *Observații asupra structurii acustice a vocalelor românești i, e, a, o și u*, *Studii și cercetări lingvistice*, XIV, 1963, 2, p. 179—198.

parlers daco-roumains.² Nous allons insister sur les nuances vocaliques spécifiques au dialecte mégléno-roumain.

Dans la notation phonétique nous avons utilisé le système de transcription phonétique de l'Atlas linguistique roumain.

Nous présentons par la suite l'inventaire des voyelles mégléno-roumaines et les valeurs moyennes des fréquences des trois premiers formants.

Voyelle	F I Hz	F II Hz	F III Hz
i	319	2 233	2 923
u	323	781	2 296
ɨ	325	1 466	2 528
e	400	2 007	2 621
á	430	1 534	2 489
o	449	991	2 387
ă	491	1 457	2 434
ɛ	519	1 971	2 570
ǔ	610	1 279	2 460
ɑ	627	1 478	2 425
a	806	1 504	2 456

LE TRAIT GRAVE-AIGU

Les voyelles [u], [o], [a], [e], [ɨ] viennent tout naturellement s'ordonner dans le mégléno-roumain, de même que dans l'aspect littéraire de la langue roumaine, depuis [u] — la voyelle la plus grave — jusqu'à [ɨ] — la voyelle la plus aiguë.

Les voyelles, appelées centrales, [a], [ă], [ɨ], communes aux deux dialectes, occupent une position intermédiaire entre les voyelles antérieures [e], [ɨ]-aiguës et les voyelles postérieures [o], [u]-graves. A côté de trois voyelles centrales existantes dans le roumain littéraire, le mégléno-roumain présente encore deux nuances vocaliques

² Voir Cornelia Coșuț, Recherches sur la structure acoustique des voyelles finales [e] et [á], *Revue roumaine de linguistique*, XI, 1966, 3, p. 268; Maria Mărdărescu, Observations sur la structure acoustique des voyelles [o] et [e] des dialectes daco-roumains, *Revue roumaine de linguistique*, XI, 1966, 3, p. 261.

intermédiaires: [ɑ] („a“fermé) et [á] (voyelle centrale à aperture réduite, placée entre [ɨ] et [ă]). Dans le cadre de la série des voyelles centrales mégléno-roumaines [a] est la voyelle la plus grave et [ɨ] est la plus aiguë. Les autres voyelles centrales [ɑ], [ă], [á] se situent graduellement entre [a] et [ɨ].

La voyelle [a] présente toutefois une différence par rapport au roumain littéraire: en effet la position de F II par rapport à F I et F III (représentée par les rapports F II/ F I et F III/ F II) est dans les deux dialectes:

$$[a] \text{ mégl. } F II/F I = 1,86; F III/F II = 1,63$$

$$[a] \text{ dr. } F II/F I = 1,82; F III/F II = 2,05$$

ce qui indique pourtant un caractère plus aigu (F II/F I > F III/ F II) dans le mégléno-roumain, correspondant du point de vue de l'articulation à une localisation antérieure de [a] mégléno-roumain par rapport à [a] daco-roumain.

Nous mentionnons que F II de la voyelle mégléno-roumaine [a] est placé dans la même zone de fréquence que F II des voyelles synthétiques [a] et [æ]³.

	F I	F II
[a] mégl.	806	1504
[a] sint.	720	1300
[æ] sint.	720	1650

La voyelle [ɛ] qui n'existe point dans l'aspect littéraire du roumain, provient généralement en mégléno-roumain de la monophthongaison de la diptongue [ea]. Antérieure comme articulation, le [ɛ] mégléno-roumain est une voyelle ouverte du point de vue acoustique.

La voyelle [ǔ]⁴ n'existant non plus dans le roumain littéraire, accuse, un caractère plus aigu que les voyelles de la série postérieure; elle est plus proche des voyelles centrales, en occupant de la sorte une position intermédiaire.

LE TRAIT COMPACT-DIFFUS

[ɨ] (F III/F I = 9,16), [ɨ] (F III/F I = 7,77) et [u] (F III/F I = 7,10) présentent dans le cadre de chaque série d'articulation la plus grande valeur du rapport F III/F I; ce sont donc les plus diffuses voyelles de la série. Etant caractérisé par la plus petite valeur de ce rapport (3,04) [a] est la voyelle la plus compacte du dialecte mégléno-roumain.

³ Pierre Delattre et colab., An Experimental Study of the Acoustic Determinants of Vowel Color: Observations on one-and two-Formant Vowels Synthesized from Spectrographic Patterns. *Word*, vol. 8, 1952, 3, p. 198.

⁴ Cf. Andrei Avram, *Asupra clasificării vocalelor românești*, *Studii și cercetări lingvistice*, VI, 1955, 3-4, p. 215-220.

Les calculs nous font voir que les autres voyelles mégléno-roumaines [e], [á], [o], [ã], [ɛ], [ø], [a] sont plus diffuses que [a], mais plus compactes que la série [i], [î], [u]. Les voyelles [a], [ø], [e], [ã] sont plus près de [a], tandis que [e], [á], [o] sont plus près des voyelles diffuses.

Les voyelles [i], [î], [u], [e], [o], [ã], [a] présentent par conséquent le même caractè-

Tableau 1

	aigu ————— grave										
diffus ↓ compact	i										
			i								
										u	
		e									
				ã							
										o	
			ɛ	ã							
								ø			
					a						
					u						

tère, par rapport au trait compact-diffus en mégléno-roumain et en roumain littéraire.⁵

La valeur du rapport F III/F I pour [â] (5,78) nous fait voir que le son est plus diffus que [ã], mais plus compact que [î]. Une voyelle centrale ayant le même degré de fermeture paraît en daco-roumain comme variante régionale et généralement en position finale, [â] (6,21).⁵

Les voyelles mégléno-roumaines peuvent être groupées du point de vue des traits grave-aigu et compact-diffus de la manière suivante:

Les matériaux examinés nous poussent à conclure que le caractère spécifique du système vocalique mégléno-roumain par rapport au daco-roumain littéraire est déterminé par les voyelles [a], [á], [ø] et [e].

⁵ Voir Cornelia Coșuț, op. cit.

DIACHRONISCHE BETRACHTUNGEN ÜBER DAS LAUTSYSTEM DER NIEDERLÄNDISCHEN SPRACHE NACH 1600

JO DAAN

Für meine Untersuchung über die Entwicklung der Vokalphoneme in der niederländischen Kultursprache in vier Jahrhunderten, habe ich einige Grammatiken und eine Beschreibung der Phonetik benützt. Letztere ist die originelle, noch heute hervorragende *Spreeckonst* von Petrus Montanus aus 1635. Die bedeutendste Grammatik ist das in seiner Art ebenfalls ausgezeichnete Werk von L. ten Kate, *Aenleiding tot de kennisse van het verhevene deel der Nederduitsche (das ist niederländische) sprake*, das in 1723 erschienen ist.

Das oberste Schema zeigt die Entwicklung der gerundeten Hinterzungenvokale; in der linken Spalte nach Montanus, in der zweiten Spalte nach ten Kate, und in der dritten Spalte, wo das niederländische Wort *nu* steht, der heutige Stand, während in der rechten Spalte, wo das Wort *straks* steht, der Zustand, den man auf Grund der herrschenden Tendenzen wahrscheinlich erwarten kann, angegeben ist. In den schmalen Spalten über denen *sp* steht, ist die übliche Schreibart der schwachgeschnittenen Vokale dargestellt, in den folgenden Spalten über denen *klank* (Laut) steht, sehen wir den vermutlichen phonetischen Wert: *zw* für schwach- und *s* für scharfgeschnitten. Die Spalte unter *d* ist nur dann benützt, wenn angegeben werden soll, daß sich ein Monophthong zu einem mehr oder weniger diphthongierten Vokal wandelte. Die phonetischen Varianten sind, wenn sie phonologisch nicht relevant sind, mit einer langen Klammer verbunden. Ein Punkt in runden Klammern bedeutet, daß die Vokallänge nicht relevant ist. Die breite Spalte unter Lambert ten Kate illustriert links den Zustand in der Kultursprache, rechts den in der Sprache der gebildeten Kreise der Stadt Amsterdam.

Da sowohl Dephonologisierung als auch Verschiebung in der Qualität und die verminderte Bedeutung der Lang-kurz-korrelation bei den Hinterzungenvokalen am deutlichsten auftraten, habe ich darauf meine Ausführungen gegründet.

Montanus unterscheidet drei Phoneme. Die beiden oberen Phoneme sind nämlich phonetische Varianten, bzw. für Zahnlaut einerseits, für Lippen- und Gaumenlaut andererseits. Die Schreibart mit *ou* ist regional und zeitgebunden, die phonetische Qualität der Variante ist aber nicht mehr nachweisbar, jedoch kann man mit Recht annehmen, daß sie geschlossener in der Lage vor Dental als vor Lippen- und Gaumenlaut war. Hundert Jahre später sind in der Kultursprache noch drei Phoneme übrig, die auch phonetisch mit drei Vokalen vertreten wurden. Zur gleichen Zeit aber kom-

men in der Sprache der gebildeten Amsterdamer, sowohl in der schwachgeschnittenen als in der scharfgeschnittenen Gruppe, nur noch zwei Phoneme vor.

Wenn man von heute noch gebräuchlichen ländlichen Dialekten, die man südlich von Amsterdam und Haarlem hört, ausgeht, ist die Annahme, daß das schwachgeschnittene O-phonem vor einerseits Zahnlaut und andererseits Gaumen- und Lippenlaut, verschiedene Realisationen kannte, berechtigt. (Auch in der Kultursprache der Gegenwart gibt es zwei Realisationen, eine in der R-lage, die andere für alle anderen Lagen.)

Aus dem Schema geht deutlich hervor, daß in der Kultursprache die Phonemenzahl in der Reihe der gerundeten Hinterzungenvokale, von drei Phonemen in 1635 und in 1723, auf zwei herabgesunken ist. Dies gilt sowohl für die heutige Kultursprache als in 1723 für die Sprache der gebildeten Kreise in Amsterdam. Hierzu soll bemerkt werden, daß was ten Kate als Kultursprache betrachtete, die Sprache war, die seiner Meinung nach in den Städten der Provinz Südholland gesprochen wurde, wobei er vor allem wahrscheinlich an die Städte Den Haag, Rotterdam und Gouda gedacht haben mag.

Ten Kate spricht einmal von Amsterdam, dann wieder von Amstelland und Rheinland. Eine ältere Grammatik des siebzehnten Jahrhunderts¹ gibt Grund zu der Annahme, daß manche Neuerungen auch im Sprachgebrauch der Stadt Haarlem bereits früh auftraten. Darum möchte ich auch lieber über Amsterdam und Haarlem sprechen. Leiden, das damals die bedeutendste Stadt im Rheinland war, hat vielleicht zu dieser Entwicklung beigetragen.

Meine *erste und dritte Behauptung*: die Dephonologisierung der Kultursprache von Amsterdam und Haarlem, möglicherweise auch von Leiden aus angefangen, habe ich mit dieser Darstellung zwar noch nicht bewiesen, wohl aber eine sehr wahrscheinliche Begründung für sie gefunden.

Meine *zweite Behauptung*: daß Verschiebungen in der Qualität der Vokale aufgetreten seien, hängt sehr wahrscheinlich mit meiner *vierten Behauptung* zusammen; daß die Bedeutung der Lang-kurz-korrelation im Verhältnis zur Qualitätskorrelation im siebzehnten Jahrhundert größer gewesen wäre als heute. Aus meinem Schema geht hervor, daß es im siebzehnten und achtzehnten Jahrhundert lange und kurze Phoneme gleicher oder fast gleicher Qualität gab. In der Sprache der Amsterdamer des achtzehnten Jahrhunderts fiel der offene gerundete Hinterzungenvokal, der in der Kultursprache noch ein Phonem war, mit dem geschlossenen Hinterzungenvokal zusammen. In der heutigen Sprache ist der geschlossenste der beiden das Phonem bei den langen gerundeten, der offene bei den kurzen Hinterzungenvokalen. Die Niederländer sind sich über den phonologischen Längenwert dieser schwachgeschnittenen Vokale noch nicht einig. Forschungen wie z. B. die von A. Cohen und seiner Mit-

¹ S. Ampzing, *Nederlandsch Taelbericht* (1628) schreibt, daß die schwachgeschnittene offene und geschlossene o nicht mehr unterschieden wurden.

arbeiter weisen wohl in die Richtung, daß in isolierten Worten die Länge relevant ist. Im Sprachgebrauch jedoch ist die Länge wahrscheinlich weniger wichtig. Die Forschungen auf diesem Gebiet sind aber noch nicht beendet.

Vergleicht man die nicht-gerundeten Vorderzungenvokale, dann steht man vor ungleich größeren Schwierigkeiten. Jedenfalls scheint im südlichen Teil der Provinz Holland der offenste Vokal noch im sechzehnten, siebzehnten und sogar noch im achtzehnten Jahrhundert mehr oder weniger palatal ausgesprochen worden zu sein. Danach herrschte über die Schreibart dieses Vokals ein hartneckiger Streit. Die meisten Grammatiker ziehen die Schreibart *ae* vor, Amsterdamer und Haarlemer schreiben aber *aa*. Die Beschreibung dieses Vokals bei Montanus läßt die Vermutung zu, daß der mit *aa* geschriebene Vokal offener als der mit *ae* geschriebene Vokal gewesen wäre. Auf Grund der Darstellung bei Montanus habe ich das *ae* bei den Vorderzungenvokalen eingeteilt.

Das Schema der Vorderzungenvokale zeigt ein unregelmäßigeres Bild als das der gerundeten Hinterzungenvokale. Im Amsterdam kam in der Sprache des achtzehnten Jahrhunderts in der Reihe der schwachgeschnittenen Vokale ein Phonem weniger vor als in der niederländischen Sprache in der Beschreibung des Montanus. Das I [y] ist nämlich diphthongiert und mit dem ursprünglichen Zwiellaut, der aus dem westgermanischen *a-i* entstanden war, zusammengefallen. Auch diese Entwicklung hat nach ten Kate ihren Ursprung in der Kultursprache von Amsterdam und Haarlem. Heute wird dieses *ei* von vielen weniger diphthongiert ausgesprochen, vor allem weil das zweite Element gesenkt wird. Obwohl dieser Prozeß noch läuft, habe ich, der Entwicklung vorgehend, das *ei* bei den Monophthongen eingeteilt. Das geschlossene *e* wird noch als Monophthong bewertet, aber die Vorliebe es zu diphthongieren, ist so auffallend, daß ich erwarte, daß es in der Zukunft unter die Diphthonge eingeordnet werden muß.

Die Zahl der scharfgeschnittenen Vokale ist von vier bei Montanus und ten Kate in der heutigen niederländischen Sprache auf drei eingeschrumpft. Auf diese Veränderungen der Vorderzungenvokale habe ich meine *fünfte Behauptung* basiert, daß nämlich die Lang-kurzkorrelation in Zukunft wieder wichtiger werden wird. Aus der Monophthongierung des *ei* wird wahrscheinlich ein Monophthong werden, das qualitativ mit dem scharfgeschnittenen *ε* übereinstimmen wird. Vorsichtshalber habe ich es aber noch mit dem umgekehrten Zeichen, das B. van den Berg² für den Diphthong gebraucht, angegeben. Außerdem werden in der Aussprache mancher Angehörigen sogenannter höheren Klassen, besonders im Westen der Provinz Süd-Holland, die Qualitäten des schwachgeschnittenen *a* und des scharfgeschnittenen *a* sich durch Verdampfung des ersten zum *a* immer ähnlicher werden. Bei meinen Ausführungen über das Schema habe ich bis jetzt ten Kates Kultursprache noch nicht behandelt. Er unterscheidet ein offenes mit Akzenten geschriebenes und ein geschlossenes *ee*.

² B. van den Berg, *Foniek van het Nederlands* (3e dr. 1964).

Das offene vertritt den westgermanischen Zwielaute *ai*, während das geschlossene aus der Dehnung der offenen Silbe des kurzen *i* oder *e* entstanden ist. Beides sind Phoneme (das offene *e* kommt auch als phonetische Variante des palatalen mehr offenen *ae* vor). Einige Forscher der niederländischen Philologie sind der Auffassung, daß ten Kate aus sprachtherapeutischen Gründen einen Zustand der Vergangenheit beschreibt und daß in der Kultursprache beide Laute, in ten Kates Zeit, phonetische Varianten eines Phonems waren. Im Westen der Provinz Holland war das wahrscheinlich der Fall; Montanus, der selber aus Delft stammt, hat sie nicht als Phoneme unterschieden. Im östlichen Teil der Provinz Südholland kann man sie auch heute noch hören und erkennt sie dort auch als Phoneme. In einer soeben erschienenen Veröffentlichung über den Dialekt der Stadt Gouda wird nachgewiesen, daß beide Laute noch zu Anfang des zwanzigsten Jahrhunderts als Phoneme unterschieden wurden.

Nun hat ten Kate zwar wahrscheinlich einen älteren Zustand beschrieben, aber in seiner Zeit kamen beide Phoneme nicht nur auf dem Lande, sondern auch in zahllosen kleineren Städten Südhollands noch vor.

Da die Wandlung der Qualität der Vokale aus den Übersichten deutlich ist, möchte ich hier nicht weiter darauf eingehen. Nun muß ich aber noch erklären, warum ich die gerundeten Vorderzungenvokale überhaupt nicht behandelt habe. Eine Einteilung dieser Vokale in das Phonemsystem wird aus ihrer Beschreibung nicht klar und deutlich. Ich habe allmählich die Überzeugung bekommen, daß es nur einige phonetische oder vielleicht sogar regionale Varianten von Vorderzungenvokalen oder von gerundeten Hinterzungenvokalen sind. Vor allem in der Lage für *r* ist die Variationsbreite sehr groß, das geht aus den Reimen auf *er* und *ar*, die im siebzehnten Jahrhundert viel vorkommen, aber auch aus den phonetischen Varianten, die man auf dem Lande hört, wo *or* sich dem *ör* nähert und *ar* und *er* wie *ör* klingt, hervor. Ten Kate beschreibt die Vokale nur vom Kieferabstand aus, den Lippenstand erwähnt er nicht, Montanus betont diesen wohl, jedoch läßt er in seiner Beschreibung die Vermutung zu, daß die Variationsbreite bei den gerundeten Vorderzungenvokalen sehr groß war. Daß die im siebzehnten Jahrhundert herrschende Unsicherheit über diesen Punkt so groß war, hängt wahrscheinlich damit zusammen, daß die Kultursprache noch im Aufbau war.

Wenn sich diese Vermutungen als richtig erweisen, wäre meine *erste Behauptung* neu zu formulieren. Die Zahl der Vokalphoneme würde dann nämlich nicht nur abgenommen haben, sondern auch die Phoneme selbst haben sich durch eine andere Verteilung über die zur Verfügung stehenden Möglichkeiten deutlicher von einander unterschieden, weil sich dann auch die Variationsbreite der phonetischen Varianten verringert hat.

Zwei Erscheinungen sind deutlich festzustellen: erstens hat im Lauf der Jahrhunderte eine Verminderung der Phonemenzahl stattgefunden und zweitens kann diese Verminderung zuerst in der Kultursprache der Stadt Amsterdam nachgewiesen werden.

Montanus 1635				Lambert ten Kate 1723									nu			straks		
				Gem. lands						Amst.								
sp.		klank		sp.		klank		klank			sp.		klank			klank		
	zw.	d.	s.		zw.	d.	s.	zw.	d.	s.		zw.	d.	s.		zw.	d.	s.
<i>ou</i>				<i>oe</i>	<i>u(.)</i>			<i>u(.)</i>			<i>oe</i>	<i>u(.)</i>			<i>u(.)</i>			
<i>oe</i>											<i>oo</i>	<i>o.</i>						<i>oɾu</i>
<i>oo</i>	<i>U.</i>		<i>U</i>	<i>oo</i>	<i>U.</i>		<i>U</i>	<i>U.</i>		<i>U</i>	<i>oo</i>	<i>U.</i>			<i>U</i>			
<i>óó</i>	<i>o.</i>		<i>o</i>	<i>óó</i>	<i>o.</i>		<i>o</i>			<i>o</i>					<i>o</i>			<i>o</i>
<i>ie</i>	<i>i(.)</i>			<i>ie</i>	<i>i(.)</i>			<i>i(.)</i>			<i>ie</i>	<i>i(.)</i>			<i>i(.)</i>			
<i>ij</i>	<i>I.</i>		<i>I</i>	<i>ij</i>	<i>I.</i>		<i>I</i>			<i>I</i>		<i>I.</i>			<i>I</i>			<i>I</i>
				<i>ee</i>	<i>e.</i>		<i>e</i>	<i>e.</i>		<i>e</i>	<i>ee</i>	<i>e.</i>						<i>e.i</i>
<i>ee</i>	<i>ε.</i>		<i>ε</i>	<i>éé</i>	<i>ε.</i>		<i>ε</i>			<i>εi</i>	<i>ε</i>	<i>ei</i>	<i>ε.</i>		<i>ε</i>	<i>ε.</i>		<i>ε</i>
<i>ae</i>	<i>æ.</i>		<i>æ</i>	<i>æ</i>	<i>æ.</i>													
			<i>a</i>	<i>aa</i>	<i>a.</i>		<i>a</i>	<i>a.</i>		<i>a</i>	<i>aa</i>	<i>a.</i>						
												<i>a.</i>			<i>a</i>	<i>a.</i>		<i>a</i>

DIE BEHANDLUNG DES STOTTERNS MITTELS ECHO-APPARAT

R. DOMANSKI, E. DZIOCH*

Die modernen Forscher interessieren sich noch immer mit der Ethologie, Symptomathologie und Therapie des Stotterns. Zahlreiche Arbeiten aus dem Gebiet und recht problematische Ergebnisse der Therapie, die Möglichkeit der Recidive und der Verschlechterung sind ein Beweis der Unvollkommenheit der Therapie. Viele moderne Autoren verbinden die Ursache des Stotterns mit dem zentralen Nervensystem, mit der Gehirnrinde und den subkortikalen Zentren, mit dem striopallidaren System. Der psychische und emotionale Zustand des Patienten sowie der Einfluß des Milieus und der zahlreichen Stimuli des modernen Lebens beeinflussen den Verlauf des Leidens.

Die Funktion der Sprache besteht aus der Tätigkeit des Atmungssystems sowie des phonetischen und artikulationischen Systems. Das leitende Impuls ist die Tätigkeit der sensorischen und motorischen Zentren der Gehirnrinde. Das Ganze richtig harmonisiert bildet außerhalb unseres Bewußtseins die fließende Sprache.

Die Regulierung dieses Sprachmechanismus ist durch das striopallidare System geleitet. Striatum und Pallidum bremsen den Muskeltonus; Striatum ist das Zentrum der autonomen Bewegungen, Pallidum ist das Zentrum für die Energie der Bewegungen und der Verminderung des Tonus. Gemäß Seeman führt die Störung der Funktion der subkortikalen Zentren zum Stottern. Schilder nennt das Stottern eine Neurose der subkortikalen Zentren.

Große Bedeutung ist den Störungen des vegetativen Nervensystems zugeschrieben. Sedláček gibt an, daß diese Störung in 80 % der Fälle des Stotterns auftritt. Fröschels weist auf die Divergenz zwischen dem Tempo des Denkens und der Sprache, was zum Stottern führt, daher die unregelmäßige Assoziation des Bewußtseins mit dem Bewegungsakt der Sprache. Gemäß Hopfer umfassen die beim Stottern aus dem zentralen Bewegungssystem der Sprachmuskeln hervorkommenden Impulse andere Muskeln, die keinen Anteil an die Sprache nehmen: das ist die Ursache der Mitbewegungen. Gewaltige affektive Reaktionen auf die äußeren Faktoren können — nach Seeman — durch Angst vor dem Sprechen manifestiert werden. Diese Logopho-

* Aus der Phoniatischen Abteilung der I. HNO-Klinik des Instituts für ärztliche Fortbildung in Warszawa.

bien verursachen Neurosen, die von Kußmaul und Gutzmann Koordinationsneurosen genannt sind. Öftere Wiederholung unregelmäßiger Impulse befestigt den unregelmäßigen pathologisch bedingten Reflex der Sprache. Dieser pathologische Punkt in der Gehirnrinde kann verbleiben und Recidive verursachen (Iwanow-Smolenski, Seeman, Tiapugin). Unsere Arbeit stellt eine der Methoden der Therapie des Stotterns mit Hilfe einer Einrichtung sog. „ECHO“ vor. Der Apparat und die Methode entstand auf Grund eines Befunds von Bernhard S. Lee im Jahre 1950, wo er die störende Wirkung des Echos auf die Geläufigkeit der Sprache wahrnahm. Diese Wahrnehmung ist in der Literatur als „Lee-Effekt“ bekannt. Ebenso ist dieses Effekt angewandt in dem Azzi-Test, der zur Entdeckung der Simulation der Taubheit dient. Ein simulierender Patient poltert und stottert bei Anwendung einer Sprachverzögerung. Nessel dagegen bewies im Jahre 1958, daß die Stotterer unter dem Einfluß des Echo viel besser sprechen.

Der von uns benutzte Apparat wurde von B. Adamczyk geschaffen und in Polen zum ersten Mal in 1958 angewandt. Die Einrichtung besteht aus dem Magnetophon mit 2 Köpfen: dem registrierenden und dem abspielenden, und aus einem Mikrofon. Der Patient spricht zu dem mit dem registrierenden Kopf verbundenen Mikrofon und der abspielende Kopf überweist den Text zum Kopfhörer. Der Patient hört also während des Sprechens seine eigene Sprache mit Verspätung wie ein Echo. Die Verspätung hängt von dem Abstand zwischen dem registrierenden und dem abspielenden Kopf und von der Geschwindigkeit des Bandes ab. In unserer Einrichtung ist die Verspätung ständig und beträgt 0.2 Sek. Wir planen eine Änderung der Einrichtung, welche eine Regulierung d. h. eine Verkürzung und Verlängerung der Verspätung ermöglicht.

Der Training mit Hilfe des „ECHO“-Apparats dient der Entwicklung der Gewohnheit des geläufigen Sprechens. Wir können das Sprechen mit Echo mit dem Sprechen im Chor vergleichen. (Adamczyk). Das Sprechen in Gegenwart der eigenen Stimme wie in der Gesellschaft einig mit dem Text vermindert die Schüchternheit und das Gefühl der Einsamkeit. Physiologisch können wir das Echo als Anregungsfaktor des Sprachzentrums mittels Gehör betrachten.

Unsere Untersuchungen umfassen 38 Patienten, die in Jahren 1965—1967 behandelt und kontrolliert wurden. Die Übungen mittels „Echo“ beginnen wir nach voller laryngo-phoniatrischer und audiometrischer Untersuchung des Patienten. In der Anfangsperiode wenden wir Psychotherapie sowie klassische Übungen des Atems und des Artikulationsrhythmus an. Die ganze Therapie umfaßt ungefähr 100 Übungen 15—20 Minuten täglich und dauert durchschnittlich 4 Monate. Zuerst üben die Patienten einzeln, dann in 2—3 Personengruppen in der Form eines spontanen Gesprächs. Außer dieser Übungen soll der Patient die Übungen zu Hause im langsamen Tempo mehrmals wiederholen. Der Lehrer in der Schule ist informiert, daß der Schüler sehr langsam sprechen soll. Der Effekt der Geläufigkeit der Sprache mit „Echo“ erscheint sofort: das setzt den Patienten in optimistische Stimmung, befreit ihn von der Spannung und erleichtert ein freies Sprechen. Anfangs dauert

der Effekt kurz, das Stottern kehrt zurück, wenn der Patient die Hörer abnimmt. Er ist aber nicht enttäuscht, weil er von dem Verlauf des Training informiert ist. Die Ergebnisse der Therapie sind in der Tafel I angegeben.

Wir haben 38 Patienten im Alter von 5 bis 51 Jahren, 5 weiblichen und 33 männlichen Geschlechts behandelt. Als „dauernde Besserung“ bezeichnen wir diejenigen

Tafel I. Die Ergebnisse der Therapie mit dem „ECHO“-Apparat.

Alter der Patienten	Verlauf der Therapie	Formen des Stotterns	Ergebnisse		
			○	◐	◑
5—10		K T T 10 7-2-1	2	4	4
11—15		12 11-1-0	1	7	4
16—20		10 5-3-2	2	4	4
21—25		3 1-2-0			3
26—30		2 1-1-0	1	1	
—51		1 0-1-0			1
	zusammen	25-10-3 38	6	19	13 38

- K — klonisches Stottern
T — tonisches Stottern
T. — tonisches Stottern mit Mitbewegungen
- Dauernde Besserung
 Besserung
 Keine Besserung
 weibliches Geschlecht
 männliches Geschlecht
 Zahl der Übungen

Fälle, in denen während 1—2jähriger Beobachtung keine Recidive vorkamen. Als „Besserung“ erkennen wir diejenigen Fälle, in welchen das Stottern seltener vorkam oder in manchen Situationen ganz verschwand, oder wenn ein Übergang von schwererer zur leichteren Form beobachtet wurde.

Dauernde Besserung wurde bei 13 Patienten erlangt, bei 19 konnte eine deutliche Besserung beobachtet werden; bei 6 Patienten konnten wir keine Besserung erlangen. Die Zahl der Übungen betrug von 5 bis 103. In manchen Fällen zeigte sich eine dauernde Besserung schon nach 1 Monat des Trainings. In anderen war eine volle Therapie von mehreren Monaten nötig. Wie man erwarten konnte, erlangte man bedeutend bessere Ergebnisse bei Kindern und bei jungen Patienten.

Dies zusammenfassend kann man diese Art der Therapie als empfehlenswert aber sicher nicht als die einzige anerkennen.

DISCUSSION

Holm:

Nach Arbeiten mit dem Logotron (Lee-Effekt) von Viennatone nach Burian sind wir heute sehr zurückhaltend mit der Anwendung des Lee-Effektes. Wir möchten hier nur die negativen Seiten aufzeigen: 1. Ein sehr sensibler Patient bekam einen leichten Nervenzusammenbruch. 2. Es trat sehr leicht Gewöhnung ein, welche die Maßnahmen wirkungslos machen. 3. Auch bei den erfolgreichen Applikationen konnten wir nach Fortnahme der Apparatur häufig ein schnelles Nachlassen des Therapieerfolges beobachten. Wir bitten um Beantwortung der Frage, wie die Erfolge und Erfahrungen mit der geschilderten Methode und unseren Erfahrungen korrelieren.

Domanski—Dziach:

1. Unsere Beobachtung betrifft nicht den portablen Logotron, sondern das stationäre Echo-Apparat, das in der Klinik installiert ist. Die Übungen sind unter der ständigen ärztlichen Kontrolle. Das erlaubt einen unerwünschten Nervenzusammenbruch zu vermeiden.
2. Nachdem die Übungen 15—20 Min täglich nicht überschreiten, haben wir keine Gewöhnung des Patienten beobachtet.
3. Unsere Beobachtung umfaßt bis 2 Jahre: die dauerhafte Besserung trat in 30% der Fälle auf. Die beschriebene Methode stellt selbstverständlich kein Panazeum dar.

DIE AUSSPRACHE DER „ɛ“-, „a“-LAUTE IN DER GEGENWÄRTIGEN POLNISCHEN SPRACHE

LEOKADIA DUKIEWICZ

In der polnischen Standardsprache sind die phonischen Korrelate der für Nasalvokale bestimmten Grapheme von der Position abhängig. Vor den Verschlusslauten bestehen sie aus dem Vokal [e] oder [ɔ] und dem Nasalkonsonanten, der im allgemeinen eine identische oder ähnliche Artikulationsstelle hat, wie der nachstehende Konsonant. Vor den Lateralen sowie vor der akustischen Pause bezeichnet man sie als Vokale mit abgeschwächter oder — besonders bei dem Laut „ɛ“ — aufgehobener Nasenresonanz.

Nur vor den Englauten treten die Laute auf, die üblich als „richtige Nasalvokale“ bezeichnet werden.

Schon Baudouin de Courtenay und später Benni haben auf die unsynchronische Aussprache dieser „richtigen Nasalvokale“ und auf die Qualität der ihnen entsprechenden, aufeinanderfolgenden Segmente aufmerksam gemacht. Ihre Beobachtungen sind im allgemeinen durch kymographische und auditorische Analysen von Koneczna und durch spektrographische Analysen von Jassem, Wierchowska und Zagórska-Brooks bestätigt worden. Die experimentalen Angaben über „ɛ“, „a“ sind jedoch nicht allzu reich und die Fortsetzung der Studien scheint begründet zu sein.

Die nachstehend beigebrachten Folgerungen sind auf Analysen gegründet, die mittels des Kay Sona-Graphs, des Tiefpaß-, Hochpaß- und Bandpaßzeitfilters, sowie auch mittels der auditorischen Methode durchgeführt wurden.

a)

Zwecks spektrographischer Analyse wurden die geprüften Laute in Logatomen „sɛs“, „sas“ von drei Personen ausgesprochen. Es wurde kein Nachdruck auf eine besonders sorgfältige Aussprache gelegt. Die Analyse wurde bei Anwendung des Breitbandpaßfilters und des HS-Frequenzganges durchgeführt.

Die erhaltenen Spektre der „ɛ“-, „a“-Laute waren deutlich zeitveränderlich. In dem ersten Teil der den analysierten Lauten entsprechenden Aufzeichnung sind die Formantfrequenzen für die Oralvokale [e], [ɔ] in der Aussprache dieser drei Personen typisch. Dann folgen Änderungen, die von dem Erscheinen beziehungsweise von der Zunahme der Nasenresonanz, sowie von den artikulatorischen Verschiebungen in dem Mundraum zeugen. Es erfolgt eine Herabsetzung der Frequenz des zweiten Formanten, die stärker für „ɛ“ als für „a“ ist. Die Spektre nehmen

Kennzeichen an, die für Hinterzungenlaute mit Nasenresonanz charakteristisch sind.

Auch die Ergebnisse von Zagórska-Brooks weisen auf die Abnahme der Frequenz des zweiten Formanten mit der Dauer der „ ϵ “- „ α “-Laute hin.

b)

Die von der spektrographischen Aufzeichnung erhaltenen Informationen wurden auditorisch nachgeprüft. Zu diesem Zwecke wurden dieselben „ ses “, „ sas “-Logatome mittels eines elektronischen Geräts zeitlich filtrierte. Die originale Tonbandaufnahme der Logatome wurde durch dieses Gerät kopiert, wobei die Grenzzeit des Tiefpaß-, Hochpaß- und Bandpaßzeitfilters je 35 ms geändert wurde. Die Dauer des Durchlaßbereiches bei Bandpaßzeitfilterung betrug 50 ms. Die auf das Tonband in einer zufälligen Reihenfolge aufgenommenen Lautausschnitte wurden von 9 Personen auditorisch beurteilt und phonetisch aufgeschrieben. Die Transkriptionen der infolge der Bandpaß- und Tiefpaßzeitfilterung erhaltenen Signale bilden folgendes Schema:

Erstes Stadium — der Oralvokal [e] oder [ɔ]. Zweites Stadium — ein Vokal derselben oder ähnlicher Qualität, wie die der oben erwähnten Laute, aber mit angelegter Nasenresonanz. Drittes Stadium — ein Laut mit zurückgezogener Artikulation und mit starker Nasenresonanz.

In den Ergebnissen der Hochpaßzeitfilterung gibt es kein erstes, orales Stadium, weil diese Filterungsart immer die nasale Endung der „ ϵ “- „ α “-Laute berücksichtigt und infolgedessen wird jeder Zeitausschnitt dieser Laute nasal gehört.

Die Diagramme, die das Prozent der als polnische Nasalvokale empfangenen Laute in Abhängigkeit von den Grenzzeiten darstellen, beweisen, daß zur richtigen Aufnahme der untersuchten Laute der Übergang von [e], [ɔ] zum hinteren, stark nasalisierten Laut von größter Bedeutung ist.

c)

Weitere Informationen über die Qualität und Quantität der den „ ϵ “- „ α “-Lauten entsprechenden Segmente haben die Transkriptionen der auf dem Tonband aufgenommenen und rückwärts auf einer Schallplatte kopierten Texte beigebracht. Die Analyse der auf solche Weise kopierten Sprachsignale sichert eine größere Objektivität bei der Beurteilung der Zeitstruktur und der Qualität mancher Laute, als eine Analyse der direkt aufgenommenen Sprache. In dem letzten Fall gibt der Hörer den orthographischen Angewohnheiten nach und beachtet nicht, daß zum Beispiel ein Laut, dem in der Schrift nur ein Buchstabe entspricht, eine polysegmentale Struktur haben kann.

Der analysierte Text bestand aus 22 Logatomen, in denen sich die „ ϵ “- „ α “-Laute sowohl vor Engelaute als auch vor Verschlusslauten befanden. Es wurde auch die sorgfältige Aussprache der untersuchten Laute vor der akustischen Pause beachtet. Man hat die [p t k]-Verschlusslaute und die [f s x]-Engelaute berücksichtigt, um die eventuellen Unterschiede in der Qualität des letzten Segments der „ ϵ “-

„ α “-Laute in Abhängigkeit von der Artikulationsstelle und von dem Artikulationsmodus des nachstehenden Konsonanten zu bestimmen.

Der volle Text wurde der Reihe nach von 16 sich der polnischen Standardsprache bedienenden Personen ausgesprochen. Rückwärts wiedergegeben, wurde er von drei Phonetikern abgehört und transkribiert. Die Analyse der phonetischen Umschrift lieferte noch einen Beweis, daß die den „ ϵ “, „ α “ entsprechenden Laute vor den Engelaute und auch vor der Pause, ähnlich wie vor den Verschlusslauten, polysegmental sind. Die Anzahl der wahrgenommenen Segmente hängt wahrscheinlich teilweise von der phonetischen Schulung der abhörenden Person ab, doch öfter wurden drei Segmente für „ α “ als für „ ϵ “ notiert. Vor der Pause nahm die Anzahl der dreisegmentalen Notierungen für beide Laute zu.

Das erste Segment wurde als der Vokal [e] oder [ɔ] identifiziert. Die Nasalresonanz wurde nicht immer beobachtet. In zweisegmentalen Strukturen folgte dann entweder der nasalierte, nichtsyllbische Hinterzungenvokal, oder ein Nasalkonsonant, oder auch ein Laut, den die Transkribierenden als einen Vokal, der sich der Nasalkonsonantenqualität nähert, bezeichneten. Vermutlich handelt es sich hier um einen [m], [n] oder [ŋ]-ähnlichen Laut, bei dessen Artikulation anstatt des Verschlusses eine Spalte im Mundraum entsteht.

In dreisegmentalen Strukturen wurde als zweites Segment immer der nasalierte, nichtsyllbische Hinterzungenvokal notiert. Ihm folgt ein Nasalkonsonant oder der vorher beschriebene Laut, dessen Qualität sich dem Nasalkonsonanten nähert.

Das letzte Segment der untersuchten Laute wurde vor den Engelaute und vor der Pause — anders als vor den Verschlusslauten — hauptsächlich vokalisch wahrgenommen.

Während vor den Verschlusslauten eine strenge Abhängigkeit der Artikulationsstelle des letzten Segments der „ ϵ “- „ α “-Laute von der Artikulationsstelle des nachstehenden Konsonanten besteht, wird vor den Engelaute häufig ein Hinterzungenvokal mit Nasenresonanz perzipiert. Vor der Pause war eine solche Abschätzung des letzten Segments die Regel.

DISCUSSION

Hamm

Die inlautenden $\epsilon : \rho$ werden von manchen als Oral-Vokale e, o + N + C aufgefaßt. Die auslautenden $-\epsilon : -\rho$ ergaben im Polnischen $-\epsilon : -\rho$ — aber: die Quantität ist nicht die gleiche und uns würde die Länge $x \frac{-ex}{-\rho}$ interessieren, die erforderlich ist, damit ein $-\rho$ noch $-\rho$ bleibt und nicht in $-o$ oder einen anderen Oralvokal übergeht. Es ist eine Frage, die nicht nur für das Polnische wichtig scheint.

НЕКОТОРЫЕ ЗАКОНОМЕРНОСТИ ВОСПРИЯТИЯ ЗВУКОВОЙ ФОРМЫ РЕЧЕВОГО ПОТОКА

З. Н. ДЖАПАРИДЗЕ*

Известно, что восприятие непривычных звуков иностранного языка представляет для человека определенную трудность. Но часто даже привычные звуки не воспринимаются человеком правильно. Это происходит тогда, когда они расположены в непривычной для слушателя последовательности. При восприятии звуковой формы как бы учитываются дистрибутивные законы языка. Так, в наших опытах аудиторы-грузины воспринимали слова родственного грузинскому сванского языка, содержащие „запрещенные“ в грузинском консонантные пары (типа *tx* в *litxe*), с искажениями. Вместо таких пар воспринимались допустимые в грузинском языке группы (типа *tx*, *tq*). Упомянутые слова сванского языка состояли из звуков, привычных для аудиторов-грузин. Непривычной была лишь их последовательность.

Во второй серии опытов тем же аудиторам предъявлялись отрезки тех же слов, но уже без „запрещенных“ пар. Сегментация производилась электронным сегментатором. Слово рассекалось на две части так, что согласные, создававшие „запрещенную“ пару, попадали в разные отрезки. Такие отрезки (типа *lit* и *xe* из слова *litxe*) воспринимались аудитором правильно или с объяснимыми отклонениями. Например, оказавшиеся после сегментации в аудиторском звуке звонкие согласные воспринимались как глухие, так как в грузинском языке звонкие согласные на исходе слова оглушаются.

Влияние дистрибутивных законов языка на восприятие звуковой формы речевого потока еще раз доказывает, что язык активно вмешивается в восприятие речи уже с первых ступеней этого процесса. Речь воспринимается не человеком вообще, а человеком — носителем определенного языка с определенным грамматическим, лексическим и звуковым строем. Безусловно, наряду с артикуляционной базой языка можно и следует говорить и об его перцептивной базе. Естественно, что эталоны, с которыми идентифицируются отрезки речевого потока при восприятии его звуковой формы, будут разными у носителей разных языков. Иначе и не может быть, так как эти

* Институт языкознания АН Грузинской ССР, Тбилиси.

эталонны представляют собой единицы, соответствующие отрезкам речевого потока различных языков с различными фонологическими системами.

В настоящее время нас интересует вопрос: каким (по величине) сегментам соответствуют эти эталоны.

Если бы эталонами являлись единицы, соответствующие отдельным звукам речи, то „запрещенные“ последовательности привычных звуков, по-видимому, должны были восприниматься правильно.

Впрочем, можно было бы предположить, что звуки из „запрещенной“ последовательности сперва воспринимаются правильно, а затем, при объединении их воспринимающим в последовательность звуков, происходит „подправка“ по дистрибутивным законам языка. В таком случае восприятие отдельных звуков речи должно оканчиваться раньше, чем восприятие любой последовательности, в которую они входят.

Однако, эксперименты показывают, что восприятие, например, начальных звуков в бессмысленном односложном звукосочетании требует большего времени, чем восприятие всего звукосочетания. (На магнитофоне воспроизводился ряд бессмысленных звукосочетаний. Аудиторам предлагалось нажатием пальца разомкнуть электрическую цепь, как только они услышат определенные звукосочетания или звуки из этих же сочетаний, переписанных на пленку несколько раз. Моменты реакций auditors регистрировались на осциллограмме. Это давало возможность измерить время от окончания записанного на той же осциллограмме интересующего нас сегмента до моментов реакций auditors. Так как условия сигнализации были одинаковы при восприятии как отдельных звуков, так и их сочетаний, разницу в длительности измеряемого периода мы считали показателем разницы во времени самого восприятия).

Полагаем, вышеизложенное говорит в пользу гипотезы, что эталонами при восприятии специфических звуковых форм в определенных условиях являются единицы, соответствующие не отдельным звукам, а допустимым в языке последовательностям звуков.

DISCUSSION

Гаприндашвили:

Докладчик убедительно доказал, что восприятие речи подчинено законам дистрибуции, характерным для конкретных языков. Не вызывает сомнения также и тот факт, что единица восприятия равняется или больше одной фонеме. Остается определить границы и объем этих единиц. Для этого, на наш взгляд, следует перед испытуемым поставить задачу выделения фонемы из более длинного отрезка речи, постепенно увеличивая объем звукосочетания.

BEITRÄGE ZUR PRÄDIKTABILITÄT DES SATZBEGINNENDEN AKZENTS IM UNGARISCHEN

LÁSZLÓ ELEKFI

Der Satz mit aktueller prädikativen Gliederung besteht mindestens aus einem Thema und einem Propositum. Unter Thema (nach der Terminologie von Charles Bally) verstehe ich das Glied, das etwas Bekanntes enthält, unter Propositum das andere Glied, welches das Novum bringt. Die Reihenfolge kann im Ungarischen Thema-Propositum oder Propositum-Thema sein. Wir wollten untersuchen, ob man in einem Satz mit aktueller Gliederung schon beim Erklingen des ersten Satzgliedes auf den Thema- bzw. Propositum-Charakter desselben Gliedes schließen kann. Wir haben Satzstücke aus etwa 150, auf Tonband aufgenommenen Sätzen ausgeschnitten. (Fünf Personen haben dieselben 30 Sätze vorgelesen, meistens Sätze aus den Gedichten von Petőfi). Unter den so erhaltenen „Satzausschnitten“ enthalten 90 ein satzbeginnendes Thema, und 70 ein satzbeginnendes Propositum. Diese 160 Satzausschnitte haben wir in gemischter Reihenfolge mit anderen Satzausschnitten zu einem neuen Tonband zusammengeklebt.

Beim Abhören waren auf Fragebogen drei Sätze als Ergänzungsmöglichkeiten bei jedem Satzstück angegeben. In jedem der drei Sätze stand der eben abzuhörende Satzausschnitt möglichst unter anderen Betonungsverhältnissen: als Thema und als Propositum. Die vierzig Abhörenden (größtenteils Studenten) sollten bezeichnen, in welchem der angegebenen Sätze das eben gehörte Satzstück wirklich stehen kann.

Die Problemstellung des Versuches hat eine phonetische und eine syntaktische Seite. Das phonetische Problem lautet: von welchen akustischen Bedingungen hängt es ab, ob wir ein Satzglied als hauptbetont (d. h. als Propositum akzentuiert) oder als zweitbetont (d. h. als Thema akzentuiert) vernehmen? Das syntaktische Problem geht die ungarische Grammatik an; hier müssen wir uns mit allgemeinen Andeutungen begnügen.

Zuerst konnten wir feststellen, daß die Mehrheit der Abhörenden nur die Hälfte der satzbeginnenden Themawörter für Thema, dagegen $\frac{2}{3}$ der Propositumwörter für Propositum aufgefaßt haben. (Unter Mehrheit verstehe ich wenigstens 51% der Abhörenden.) Dieses Ergebnis bedeutet, daß der Propositumakzent des satzbeginnenden Wortes mit größerer Sicherheit erkennbar ist, als der satzbeginnende Themaakzent. Das hängt mit dem hervorhebenden, oft emphatischen Charakter des Propositumakzentes zusammen, während der Themaakzent nur selten einen besonders geprägten Ablauf hat.

Betrachten wir nun einen Fall, wo der Themaakzent besonders ausdrucksvoll geprägt war: *Az idő* (Die Zeit) aus dem Satze *Az idő akkőzben haladott sietve* (Die Zeit verging inzwischen eilends). Alle Abhörer haben dieses Satzstück (d. h. das Substantivum mit dem Artikel) als Thema eines Satzes apperzipiert. Die Schalldruckkurve zeigt eine verhältnismäßig größere Dauer und einen höheren Pegelwert der akzentuierten, aber phonologisch kurzen Silbe *i*. Die Tonhöhe dieser Silbe ist freilich höher (ungefähr um eine Quarte) als die des Artikels, beide Silben sind steigend, die nachfolgende Silbe ist etwas niedriger.

Ein anderes Beispiel stellt einen ausdrucksvoll geprägten Propositumakzent dar: *Napfogyatkozás* (Sonnenfinsternis) aus dem Satz *Napfogyatkozás jött szeme világra, melynek elmúlását hasztalanul várta* (Sonnenfinsternis kam auf sein Augenlicht, auf deren Vergehen er umsonst wartete). Der Schalldruckwert erhebt die Akzentsilbe über die anderen Akzentsilben im Hauptsatz; dazu gesellt sich der melodische Faktor: nach Anstieg im Konsonanten *n* eine kaum steigend-fallende Melodie im Vokal *a*, dann ein Fall um mehr als eine Quarte und ein weiteres allmähliches Sinken in den folgenden vier unbetonten Silben. Also nicht der Schalldruck allein, sondern mehrere Faktoren: Dauer und Tempoänderungen, Lautdruck und Melodieverlauf zusammen sind unverkennbare Merkmale des Propositumakzents.

Folgende Beispiele sollen solche satzbeginnende Ausschnitte darstellen, deren Akzentcharakter nicht leicht erkennbar ist.

Zunächst ein nicht gut ausgeprägtes Thema: *Szemeid* (deine Augen) aus dem Satze *Szemeid gyöngék . . . a napfény megvakít* (Deine Augen sind schwach; der Sonnenschein wird dich blenden). Nun ein einsilbiges Thema eines Ausrufsatzes, zuerst mit ausdrucksvoller, dann in einer weniger ausdrücklichen Vortragsweise. . . Nun folgt das Propositum *napfogyatkozás* (Sonnenfinsternis) in zwei anderen Sprechvarianten; die nächste Variante war nicht als Propositum erkennbar . . . Und was jetzt folgt, war statt Propositum größtenteils als Thema beurteilt. Dann ein Thema, welches vollständig als Propositum aufgefaßt worden ist: *a napfény* (der Sonnenschein): Schließlich hören Sie, bitte, noch ein einsilbiges Beispiel an, ein nicht gut geprägtes einsilbiges Propositum: *ez* (das) aus dem Satze *ez volt végső lehellete: Jancsikám* (das war ihr letzter Atem: Hänschen, mein Hänschen).

Von den syntaktischen Beziehungen unserer Untersuchung möchte ich folgendes herausheben. Unter den satzbeginnenden Wörtern der 160 Sätze können wir drei morphologische Typen unterscheiden: 1. Substantive mit dem bestimmten Artikel, 2. Substantive ohne Artikel, 3. Adverbia und Pronomina. Die Substantive mit bestimmtem Artikel, in denen die Determiniertheit auch formell bezeichnet ist, haben größtenteils Themacharakter, die Hälfte der Substantive ohne Artikel spielen die Rolle des Propositums, etwas mehr als die Hälfte der Adverbia und Pronomina kommen schließlich in der Rolle des Themas vor. Bei den Substantiven mit dem bestimmten Artikel ist der Akzentcharakter beider Art verhältnismäßig am besten erkennbar, bei den Substantiven ohne Artikel ist der Akzentcharakter nicht erkennbar, bei Adverbiis und Pronomina ist das Propositumakzent gut erkennbar, der Themaakzent aber

nicht. In den Typen, wo der Akzentcharakter nicht erkennbar ist, mag er wahrscheinlich auch nicht relevant sein. Das Problem des Erkennens beim Hören können wir also umkehren: unterscheidet der Sprecher die charakteristischen Akzentweisen im Ungarischen? Nicht immer. Die Sprecher haben die Akzentweisen dort miteinander vermischt, wo die Verständlichkeit der Sätze nicht auf der feinen Prägung des Akzents, sondern auf der lexikalisch-grammatischen Konstruktion des Satzes beruhte.

Noch etwas von den silbenprosodischen Bedingungen. Bei den Hauptwörtern mit dem bestimmten Artikel zeigt unsere Untersuchung einen Zusammenhang mit der Silbenzahl: je länger das Substantiv ist, desto mehr kann man seinen Akzentcharakter erkennen. Bei den Hauptwörtern ohne Artikel, sowie bei Adverbiis und Pronomina läßt sich kein Zusammenhang zwischen der Silbenzahl und dem Erkennen des Akzentcharakters feststellen. Der bestimmte Artikel macht also den Sprecher aufmerksam, daß er beim folgenden Substantiv besser auf die Akzentuierung achten soll. Dadurch hängt die logische Funktion des Artikels mit dem Mitteilungswert des Satzgliedes zusammen.

DISCUSSION

Beneš:

In dem interessanten Versuch wurden die Wechselbeziehungen zwischen Thema-Rhema-Gliederung, Syntax und Klanggestalt eines Satzes untersucht. Sollte man nicht auch die Satzsemantik ins Auge fassen?

Daems:

Bei diesem interessanten Experiment ist es absolut notwendig zu wissen, wieviel Personen die Sätze auf das Band gesprochen haben, weil man sonst leicht Gefahr liefe aus der Aussprache einer zu kleinen Zahl von Personen allgemeine Konsequenzen zu ziehen. Um generalisieren zu können soll man unbedingt eine große Zahl von Sprechern und Informanten verwenden.

Lehiste:

Ich nehme an, daß fast alle Ungarn die gesamten Gedichte von Petöfi auswendig kennen. Dürfte es nicht sein, daß die Versuchspersonen die Gedichte einfach erkannt haben?

Skalička:

Gibt es einen Unterschied zwischen den dichterischen und anderen Texten?

Elekfi:

Ad Beneš: Die Betrachtung der Satzsemantik ist in diesen Untersuchungen äußerst wichtig. Aber in der ersten Phase der Untersuchung begrenzte ich mich auf Satztypen, die mit Substantiven (eventuell mit dem bestimmten Artikel) oder mit Pronomina oder Adverbiis beginnen. Ein weiteres Ziel der Untersuchung ist eben die Festlegung der Kriterien, nach denen ein Satz auch kontextfrei eindeutig ist. Im Ungarischen hängt das meistens nicht von der Bedeutung der Wörter, sondern von ihrer Folgenreihe und Konstruktion im Satze ab.

Ad Lehiste: Viele Gedichte von Petöfi sind nicht allgemein bekannt. Nur wenige unserer Abhörsätze waren den bekanntesten Gedichten entnommen. Andere sind teilweise aus weniger bekannten Gedichten, teilweise aus Prosawerken von Petöfi, teilweise selbstkonstruierte Sätze. Und die aus Gedichten ausgeschnittenen Satzanfänge kommen meistens auch in anderen Sätzen vor, so daß die Wahrscheinlichkeit des Erkennens minimal war und auch in diesen Fällen oft auf der Betonungsweise beruhte.

Ad Skalička: Augenblicklich kann ich keine entscheidende Antwort geben. Die Frage kann erst nach der weiteren Bearbeitung des Untersuchungsmaterials beantwortet werden.

DIE ENTWICKLUNG EINES TONHÖHENSCHREIBERS FÜR PHONETISCHE FORSCHUNGEN

HEINRICH ERAS*

Im Rahmen der Beschäftigung mit der Satzintonation im Hochdeutschen und in den deutschen Mundarten wurde die Aufgabe gestellt, einen Tonhöhenschreiber aufzubauen. Das Gerät ist ausschließlich für phonetische Forschungen bestimmt und soll bei vertretbarem Aufwand einen großen Frequenzbereich hinreichend sicher und genau analysieren. Es besteht nicht die Notwendigkeit, eine möglichst kleine Zeitverzögerung zwischen Eingangs- und Ausgangssignal zu realisieren, wie sie für Analyse-Synthese-Geräte (z. B. Vocoder) gefordert werden muß.

Zunächst wurden einige Schaltungen und Verfahren auf ihre Verwendbarkeit hin analysiert. Nach der Klassifikation von Ungeheuer (1) schieden im Zeitbereich arbeitende Schaltungen wegen ihres hohen Aufwandes (2) oder wegen ihres begrenzten Frequenzumfanges (3) aus. Schaltungen mit Impulsgerichtung (4) schienen wegen ihrer großen Zeitkonstante der Anzeige nicht geeignet. Ebenso sollten Parallelschaltungen von Teilsystemen mit eingegängtem Arbeitsbereich (5) nach Möglichkeit vermieden werden.

Von allen Verfahren erwies sich für die gestellte Aufgabe eine im Frequenzbereich arbeitende Filterschaltung als günstigster Kompromiß zwischen Aufwand und erreichbaren Ergebnissen. Es wurde im Prinzip das bekannte Verfahren von Grütz-macher-Lottemoser (6) gewählt. Die zur Anzeige verwendete Impulszählschaltung (getriggert Sägezahn) ergibt ein gutes optisches Bild ohne Zeitkonstanten und gestattet, Fehlanzeigen relativ leicht zu erkennen. Durch zusätzliche Stufen wird versucht, einen großen Frequenzbereich ohne Umschaltung, einen großen Dynamikbereich sowie eine geringe Beeinflussung durch Brummen und Rauschen zu erreichen.

Anhand des Blockschaltbildes (Abb. 1) wird im folgenden das aufgebaute Gerät beschrieben.

Der regelbare *Vorverstärker* dient zur Anpassung an alle gebräuchlichen Spannungsquellen (Mikrophon, Tonbandgerät etc.).

Es folgt ein *Verzerrer* zur Verstärkung bzw. Regenerierung des Grundtones. Es wird ein Spannungsteiler aus je einem Widerstand und zwei antiparallel geschalteten

* Deutsche Akademie der Wissenschaften zu Berlin, DDR.

Dioden im Längs- und Querzweig verwendet. Der Verzerrer erwies sich bei Versuchen anderen aus der Literatur bekannten Schaltungen (5, 6, 7) überlegen.

Der *Bandpaß* gestattet eine grobe Vorauswahl des Analysierbereiches. Es wird eine elektronische Filterschaltung mit einer Flankensteilheit von 12 dB/Okt., minimalem Überspringen und kontinuierlich regelbaren Grenzfrequenzen verwendet.

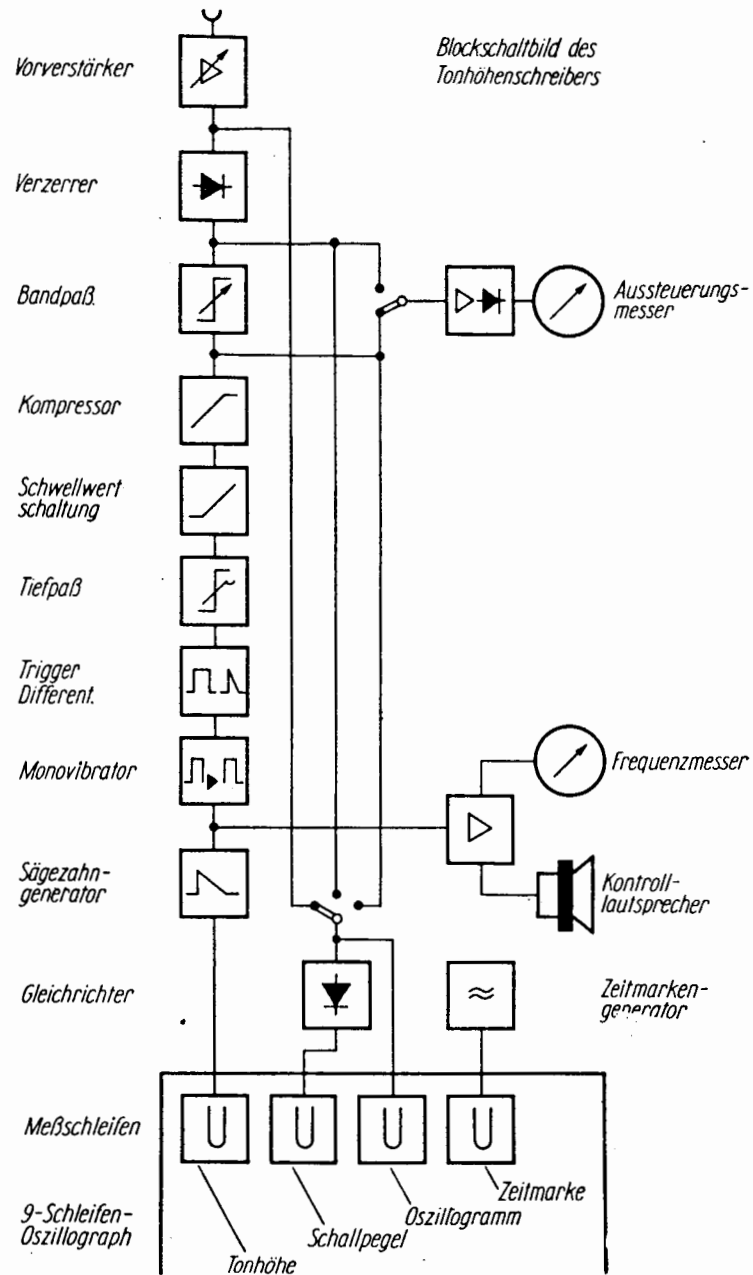


Abb. 1.

Die *Aussteuerungsanzeige* erfolgt normalerweise hinter dem Bandpaß, so daß starke Formanten und Geräuschanteile die Anzeige nicht beeinflussen können. Zur Verwendung gelangt ein Spitzenwertmesser mit logarithmischer Anzeige zwischen -46 dB und $+6$ dB. Die Einschwingzeit beträgt < 10 ms, die Abklingzeit $> 1,5$ s. Eine genaue Aussteuerung und Spitzenwertanzeige ist notwendig, um alle Stufen des Tonhöhenmessers optimal zu betreiben.

Der auf den Bandpaß folgende *Kompressor* erweitert den Dynamikbereich nach oben hin um ca. 18 dB, ohne den Störabstand bei niedrigen Pegeln zu verschlechtern. Die Einregelzeit beträgt ca. 1 ms, die Ausregelzeit ist mit ca. 50 ms bewußt niedrig gehalten, so daß auf hohe Pegel folgende schwache Signale nicht beeinflußt werden können. Die Regelschaltung arbeitet mit einer gesteuerten Diodenbrücke in Vorwärtsregelung.

Die *Schwellwertschaltung* ist erst für Signale über -35 dB (bezogen auf Vollaussteuerung) voll durchlässig und gestattet so eine Unterdrückung restlicher Störspannungen.

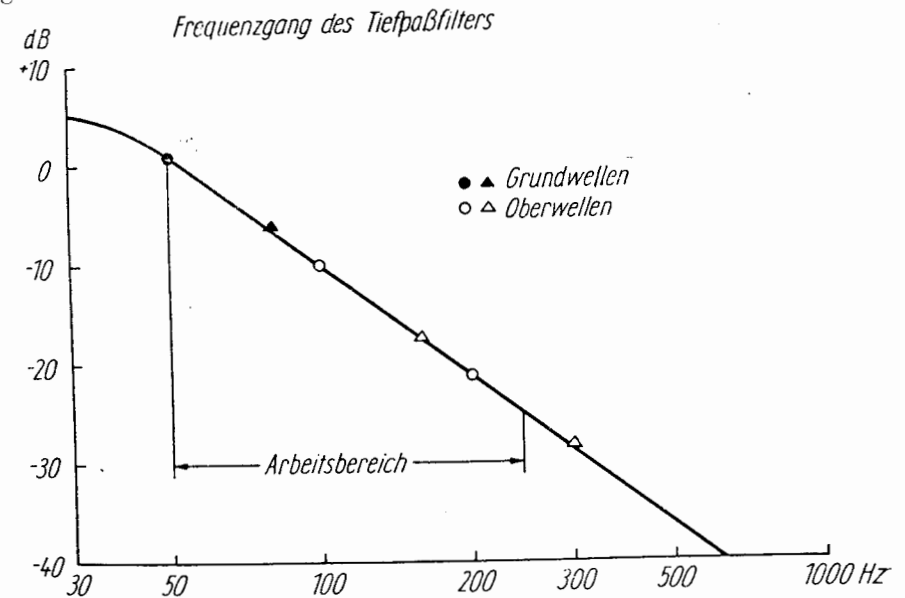


Abb. 2.

Entscheidend zur Erweiterung des Frequenzbereiches trägt das folgende *Tiefpaßfilter*: (Abb. 2) bei. Es handelt sich um ein zweistufiges, entkoppeltes RC-Filter mit ca. 12 dB Flankensteilheit, dessen Grenzfrequenz stets unter der tiefsten Signalfrequenz liegen soll, so daß immer auf der abfallenden Flanke gearbeitet wird. Damit wird die erste Oberwelle im Verhältnis zur Grundwelle stets um 12 dB, die zweite um 24 dB usw. gedämpft. Das Filter ist stufenweise für verschiedene Stimmlagen umschaltbar.

Die Ansprechschwelle des *Schmitt-Triggers* muß so niedrig liegen, daß seine An-

steuerung auch bei hohen Grundfrequenzen, die durch das Tiefpaßfilter natürlich auch geschwächt werden, noch gewährleistet ist. Der Trigger formt das annähernd sinusförmige Eingangssignal in ein Rechtecksignal mit konstanter Amplitude um. Durch anschließende Differenzierung und Gleichrichtung wird pro Grundfrequenzperiode ein positiver Impuls gewonnen. Diese Nadelimpulse werden im *Monovibrator* auf eine konstante Breite von 0,6 ms gebracht. Die Impulsfolgefrequenz ist der Periodendauer des Eingangssignales gleich.

Der *Frequenzmesser* gestattet eine Kontrolle und Abschätzung der zu messenden Tonhöhen. Das isolierte Tonhöhenignal kann auch mit einem kleinen *Kontrolllautsprecher* abgehört werden.

Prinzipschaltbild des Sägezahngenerators

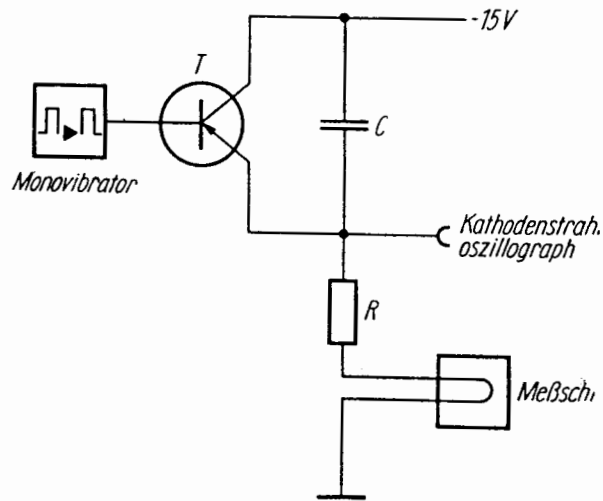


Abb. 3.

Der *Sägezahngenerator* (Abb. 3) wandelt die Impulse des Monovibrators in Sägezahnimpulse um, deren Anstiegszeit der Periodendauer proportional ist. Damit bildet die Amplitude des Sägezahnes ein Maß für die Periodendauer. Während der Grundfrequenzperiode lädt sich der Kondensator C über den Widerstand R auf; der Transistor T ist gesperrt. Während des 0,6 ms Impulses vom Monovibrator wird der Transistor durchgesteuert und entlädt den Kondensator C sehr schnell. Trotz ihrer Einfachheit besitzt diese Schaltung eine hohe Stabilität, da die Aufladung von C in einem weiten Bereich nicht von den Eigenschaften des Transistors T beeinflusst wird. Ein weiterer Vorteil ist, daß am Ausgang keine hohe Gleichspannung liegt, so daß die jeweiligen Schreiber direkt angekoppelt werden können. Durch die nicht-lineare Aufladung von C ergibt sich eine annähernd logarithmische Frequenzanzeige. Der Ausgang wurde an einen 9-Schleifen-Oszillografen angepaßt. Ferner ist ein Ausgang für einen Sichtspeicheroszillografen vorgesehen.

Alle Betriebsspannungen werden einem stabilisierten Netzteil entnommen. Der Aufbau erfolgt bausteinweise auf Leiterplatten mit gedruckter Schaltung. Die einzelnen Baugruppen sind durch steckbare Kontaktleisten verbunden. Der ausschließliche Einsatz von Halbleiterbauelementen wirkt sich besonders in den Impulsschaltungen, im Kompressor, Verzerrer und in der Schwellwertschaltung günstig aus. Das ganze Gerät ist in ein 19" Einschubchassis eingebaut.

Die ersten praktischen Versuche mit dem Gerät haben gezeigt, daß es den gestellten Anforderungen gerecht wird. Da in der Aufzeichnung außerdem Pegel und Oszillogramm (Abb. 4). mitgeschrieben werden, kann das Gerät auch zur Segmentation und zur Kontrolle von Transkriptionen herangezogen werden. Um eine noch höhere Sicherheit in der Anzeige zu erhalten, wird gegenwärtig ein neues Verfahren entwickelt, welches im Zeitbereich und ohne Filter arbeitet.

LITERATUR

1. G. Ungeheuer: Zur Periodizitätsanalyse phonetischer Signale. *Phonetica* 10, 1963.
2. L. O. Dolanský: An instantaneous pitch-period indicator. *JASA* 27, 1, 1955.
3. W. Tscheschner: Ein Beitrag zur Messung der Stimmbandgrundfrequenz. III. Akustische Konferenz in Budapest 1964.
4. F. Winkel: Entwicklung eines Tonhöhenmessers für Sprachuntersuchungen. Speech Comm. Seminar, Stockholm 1962.
5. A. Risberg: Fundamental frequency tracking. *Proc. 4th Int. Congr. Phonet. Sc.* 1962.
6. Grützmacher, Lottemoser: Über ein Verfahren zur trägheitsfreien Aufzeichnung von Melodiekurven. *Akustische Zeitschrift* 2, 1937.
7. Carré, Lancia, Paillé, Gsell: Étude et réalisation d'un détecteur de mélodie pour analyse de la parole. *L'Onde Electrique* 434, 1963.

DISCUSSION

Froekjaer—Jensen:

1. Do you know the signal-to-noise ratio for high fundamental frequencies?
I am asking this because the intonation curve which, as you just have shown, possesses some of these irregularities that we in Copenhagen have found typical for high pitched intonations.
2. I do not believe that your LP-filter with a slope of 25 dB per octave is the best one for extracting the fundamental frequency. Referring to different investigations at the Speech Transmission Laboratory in Stockholm and the investigations of my own, the best filter slope seems to be that of 18—20 dB per octave. For most subjects this filter gives the smallest amount of errors. If the sloping of the LP-filter is smaller than 18 dB/oct. you may be registering the second harmonics instead of the fundamental frequency, and if the sloping is greater than 20 dB oct. you will get the above mentioned uneven intonation line, caused by a damping of the fundamental frequency.

Kahl:

Muß der Durchlaßbereich des verwendeten Bandpasses der Sprachgrundfrequenzlage des Sprechers (männlich bzw. weiblich) angepaßt werden? Ist absolute oder relative Bandbreite konstant oder keine von beiden?

Smith:

In case you want to have still better intonation curves I should propose that at a later time we get together in Prague and take this opportunity to assemble all specialists in the construction of fundamental frequency meters.

Eras:

Bisher zeigte das Gerät bei tiefen Frequenzen keine Fehlanzeigen.

Jedes System hat spezifische Zeitkonstanten. Es wurde lediglich versucht, in der Anzeige schaltung zusätzliche Zeitkonstanten zu vermeiden.

Schwankungen der Anzeige sind wahrscheinlich auf geringen Störabstand des Tonbandmaterials zurückzuführen.

Durch Anwendung von 2 Filtern (Bandpaß- und Tiefpaßfilter) ergibt sich eine resultierende Flankensteilheit von ca. 24 dB/Okt.

Eras: Die Entwicklung eines Tonhöhenschreibers für phonetische Forschungen

zeitmarke 0,1 s

oszillogramm

schallpegelkurve

tonhöhenkurve

Hz
400
200
100
50

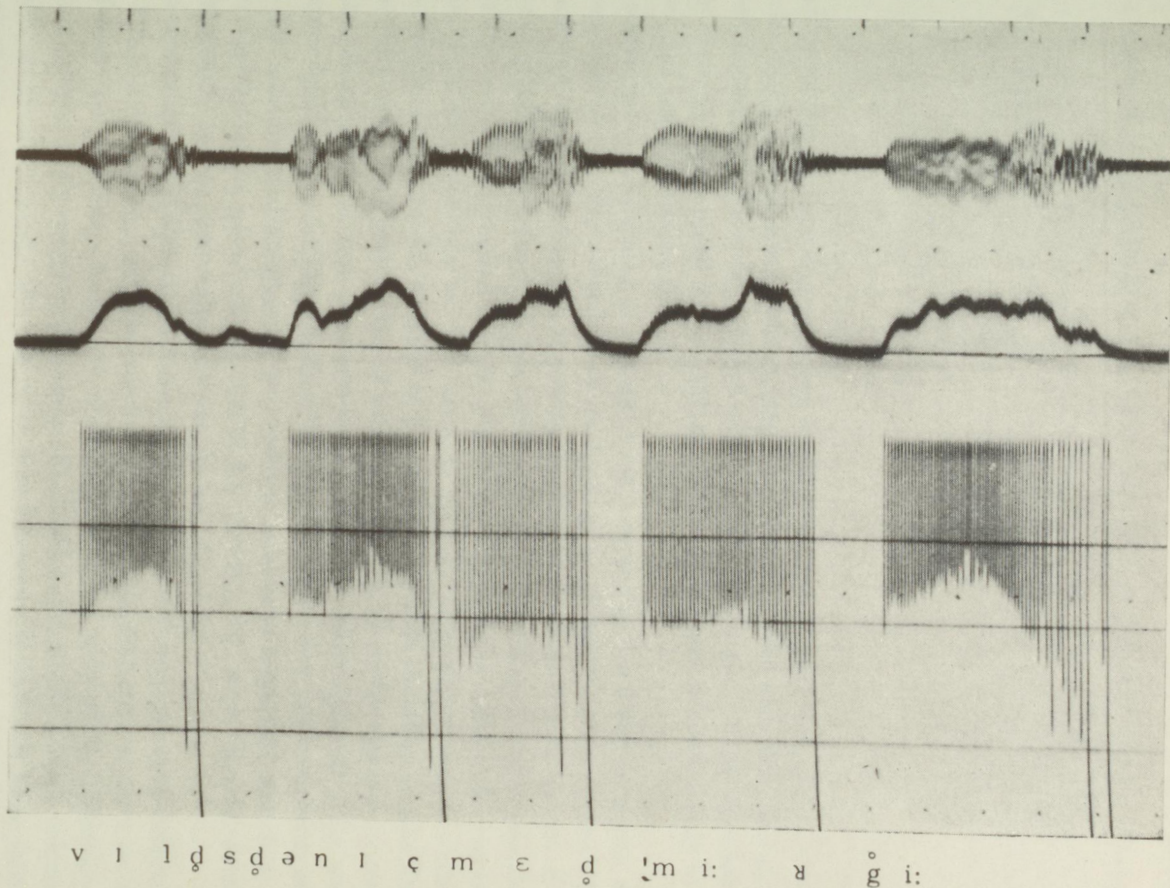


Abb. 4.

PRESENT CHANGES IN ENGLISH PRONUNCIATION

S. S. EUSTACE

This paper results from a comparison between my speech and that of five 17-year-old schoolboys now at Eton College, and born between 1949 and 1951. We all live in west London and have very similar origins, and I was myself at the same school. We are thus environmentally similar, and our differences of speech are due only to the twenty-odd years difference in age.

The material is a recording of 170 sentences, containing 470 different words, read by them and by me. The reliability of the findings is limited: ((Firstly, inasmuch as they are derived from subjective impressions; the statistics are merely counts of the written symbols in the transcript, and terms such as lip rounding must be understood as relating, not to the original sounds, but to my attempts to imitate the recorded sounds. Secondly, inasmuch as they are true only for a particular social context, where a single informant was closeted with a machine and an investigator and knew neither very well. The reservation of social context is of course implied in every linguistic description. In free discourse both I and the informants show far wider departures from the kind of English described by Pr Jones. For instance: from each of the five informants; *Sunningdale -deo*; *railway* ¹reowe; *can't one* ¹kãʔwan; *kick the bucket* ¹kɪʔ ðə ¹bakɪʔ; *about* ¹bɔt *twenty-four* ¹twɛɪʔfɔ; or from myself, *twenty-three* ¹twen¹θri, *isn't it* ¹ɪnɪʔ, *what will that be* ¹wɒl ¹lɛʔbi.))

I was concerned only with the differences between me and them. In each difference noted I therefore consider only the word in which it occurred once or more on the tape. I state the percentage frequency of occurrence of one term of the difference as against the other, in my speech and in theirs. For instance, in the case of the velarised [ɫ], which occurs in 92 words, the difference is as between a lateral and a non-lateral or sulcal; that is, with the tongue touching or not touching the palate. The 92 words gave to me 92 occurrences of velarised [ɫ], of which 98 % were lateral; and to them, 92 multiplied of course by 5, that is, 460 occurrences, of which 46 % were lateral.

In the findings themselves, we note an increased occurrence of glottalised consonants and [ʔ].

In the case of [p], mine was glottalised in 36 % of cases, theirs in 54 %.

In the case of [t], and [ʔ] replacing [t]: for me 55 %, for them 70 %.

In the case of [k], and the [ʔ] here and there replacing [k]: for me 44 %, for them 61 %.

There was also glottalisation of [tʃ] after a stressed vowel. Of these cases, I had glottalisation in 67 %, and they had it in 89 %.

[ʔ] also occurred in other situations; for instance, instead of linking [r], as in *joʔl æp:lɪke:n*; or between vowels, as in *ði ʔaɪlənd*; or to replace an initial [ə], as, *ʔfeɪməs lænma:k*; or between a nasal and a voiceless fricative, as, *ɪkʰs pænʔn* or *leɪʔθ*. In such cases, I had it in 11 % of all the words concerned, but they had it in 33 %.

To test for assimilation of [d] or [n] to the corresponding velar or labial, I included some words, like *brown-coal*, *skinfint* or *threadbare*, which would probably show assimilation in substandard speech. But there were few examples of [d] assimilated except after [l], as, *lɪgəlbma:n* or *lɪ:lgkra:ft*. Except in very common words all the bilabial assimilations of [db], such as *lɔrebbeə* for *lθredbeə*, were the work of a single informant.

[n] was assimilated more often, in about 30 of the sentences. But there were also about 15 candidates, such as *mainmast* or *organ-pipes*, which evoked no assimilation. Most of the [n] assimilations, like *lɪfɪlɪ:ɪ:ɪ:gɪnɪz*, were velar. The figures for [n] assimilated were 18 % for me, 47 % for them.

Among the fricatives we note the assimilation of [v] to [β], as, *the rights of man-tsaβlɪm*. This occurs very predictably with the informants. The figures for [v] assimilated to [β], [b], [f], [z], and so on, were 24 % for me and 56 % for them.

There is regressive assimilation of the voiced dental fricative [ð] after [n], [l], [z] and so on in phrases like *lɪnnəbɒks*, *lɪlləgæp*, *lɪvʔszəlɪtarm*. The percentages for this assimilation were 17 % for me, 42 % for them.

As for the flap [ɾ]: after [θ] as in *θɪəu*, 100 % for me, 67 % for them; between vowels, as in *veɪu*, 48 % for me, 5 % for them.

The recording of me had no instances of [r]'s with lip-rounding, retroflexion, or what I shall call the lingua-faucal quality which distinguishes the so-called bunched [r]. The percentage of these sounds is accordingly 0 % for me, but 23 % for the informants.

We also note the change of velarised [l] from a lateral to a non-lateral sound. My velarised [l]'s were 98 % lateral, but theirs, only 46 %. The remainder of theirs were roughly in the proportions sulcal 7, vocalic [u] etc 2, entirely absent 1, depending partly on context.

The number of long consonants has also grown.

The post-war generation show a marked lowering of [æ] towards [a], particularly the girls. In this survey there was appreciable lowering in 6 % of my words, against 41 % of the informants'. Before velarised [l], three of the five made either no distinction, or sometimes a distinction of length only, between [æ] and [ʌ] unstressed, having exactly the same vowel in *Algeria* and *ulterior*.

[u:] may be fronted to [u:], or diphthonged to [ɪu], or to a diphthong whose first element is unrounded, [ɪu]. If we group on the one hand all those occurrences which

contain either fronted or unrounded sounds, and on the other, those which are not thus altered, then the altered ones are 41 % of mine but 71 % of theirs.

[eɪ] and [aɪ] showed little difference, except before velarised [l], where they tended to be "smoothed" to [e:] and [a:]. Thus *neɪl* might become *ne:u* and *waɪld*, *wa:lɪd*. This smoothing happened with me in 13 % of cases, but with them, in 53 %.

The second element of [ɔɪ] before velarised [l] was rounded and retracted, giving [ɔɪʊ]. The word *foiled* sounded like my *furled*. This rounding or retraction happened in 25 % of cases with me, and in 53 % with them.

The diphthong [aɪ] is largely unrounded to [aɪ], except before velarised [l]. The change is widespread, so the difference between me and them is less marked. My [aɪ] was unrounded in 60 % of cases, and theirs in 77 %.

The [əu] in free position is also largely unrounded to [əu]: for me, 39 %, for them, 84 %. Of their rounded ones, 72 % were before a lip consonant. Before another vowel, the diphthong is levelled to [ɛ:] as when *pəʊətri* becomes *pɛ:ətri*, or *fləʊɪŋ flɛ:ɪŋ*. This levelling happened in 40 % of cases with me, and in 56 % with them.

I was surprised that [əu] before velarised [l] was sometimes rendered [əu]: *told* was sometimes pronounced *təʊld*, with no [l] at all. I would certainly have called this a vulgarity, yet there were several instances of it. One informant made no difference between *coal* and *col*, but had *khəu* for both. But another pronounced *pole* and *pearl* identically.

There were formal differences in about 60 words, including spelling-pronunciations like *ɪkʰs æmənd* for *ɪgʰz æmənd*, or vulgarisms like *θɪlɛtə* for *lɔlɛtə*. The spelling pronunciations and vulgarisms accounted for 20 % of the formal differences with them, against 3 % with me.

Time forbids discussion of biographical influences on the speech of the informants; of the phonetic context of the changes; of aspiration, other glides, intonation, the proportion of diphthongs to simple vowels; or of the phonology of each idiolect. Further interesting questions would also certainly emerge from spectrographic study. But I hope to deal with some of these points in a future report.

Although the new features often seem to resemble Cockney, their origin is rather to be sought in the English of the middle classes, a vast but ill-documented dialect with which the informants have had an increased contact; as a formative influence, the governesses of a former generation have now vanished, and are replaced by the somewhat wider social range of the infant school.

DISCUSSION

Schubiger:

It was said in the summary that no general trends in stress were observed as e.g. 'educational', 'universal' etc. Did the test include enough material in which a change of stress might have become evident?

Lewis:

The speaker was asked whether he had noted in the cases of phonetic variation any which might be attributed to the sort of paralinguistic differences that might be reflection of the changing mood, increasing fatigue etc of his subjects (e.g. in their choice of flapped *r* as opposed to other types).

Sivertsen:

Mr. Eustace, you gave precise figures for the occurrence of such features as lip-rounding, glottalization, vocoid replacement for the lateral ("dark l"), etc. How did you observe the occurrence of such features what were your means of observation?

Eustace:

ad Schubiger: My remark in the summary was erroneous. There were 11 words in which —'— > '—'— was possible. I had —'— in all, but the informants had '—'— as follows: *Algeria* 0%, *co-operative* 0%, *Be quiet* 40%, *malpractices* 40%, *Minoan* 56%, *Mulholland* 40%, *Nairobi* 40%, *Niagara* 25%, *Nigeria* 20%, *ulterior* 0%, *Victoria* 40%. In this context there is evidently a tendency for stress to become a function of vowel quality.

13 other words with stress difference were too diverse to allow a firm conclusion.

ad Lewis and Sivertsen: An answer is now included in double brackets in the text.

SPRACHNORM UND SPRECHNORM

KENOSUKE EZAWA

Der besondere, normative Charakter der Sprache bringt es mit sich, daß man ihr als Gegenstand nicht in der sprachlichen Äußerung selbst, sondern erst in deren Bezug auf eine Norm begegnet. Diese Erkenntnis hat sich heute allgemein durchgesetzt, worauf die verschiedenen methodischen Ansätze strukturell-linguistischer Richtung mit ihren charakteristischen Begriffsbildungen, welche sich jeweils auf die Sprachnorm und auf die Sprachäußerung beziehen, eindeutig hinweisen. Nur wird der Begriff der Norm vielfach gleichsam naturalistisch, d.h. determinativ und eindimensional, aufgefaßt, weshalb manchmal der an sich fruchtbare methodische Ansatz in einem flachen Formalismus, in dem man den Gegenstand Sprache selbst nicht in den Griff bekommt, enden muß.¹ Heute scheint es mehr denn je nötig zu sein, sich mit dem Begriff der sprachlichen Norm ernstlich zu befassen.

Im Fall des hochdeutschen Lautsystems sollte man vor allem die weitgehende funktionelle Labilität mancher geltender Lautnormen von vornherein ins Auge fassen und im Zusammenhang damit die besonders feste Bindung dieser Normen an die hochdeutsche Schriftsprache, d.h. geschriebene hochdeutsche Sprache, mit ihrer ausgeprägten Orthographie berücksichtigen. Es ist überhaupt in Frage zu stellen, daß das hochdeutsche Lautsystem ein in sich geschlossenes, im wörtlichen Sinne eindeutiges normatives System darstellt, dem man auf Grund rein formaler Analysen beikommen könnte. Es sollte hier vielmehr neben dem rein funktionellen ein anderer Normbegriff zur Geltung gebracht werden, und zwar ein solcher, der die normative Abhängigkeit der lautlichen Äußerungen von schriftlichen Fixierungsformen der Sprache, die über die funktionellen Normen hinweg besteht, impliziert. Diese Möglich-

¹ In diesem Zusammenhang sei besonders verwiesen auf Eugenio Coseriu in "*Sistema, norma y habla*" (Revista de la facultad de humanidades y ciencias, Año 6, No. 8 Universidad de la Republica, Montevideo 1952, S. 113—181): „Vom Standpunkt des Sprachsystems aus können wir nun die Normen und das konkrete Sprechen als aufeinanderfolgende Stufen der Realisation desselben betrachten. So gesehen, erscheint das Sprachsystem als ein System von Möglichkeiten, von Koordinaten, welche die offenen und versperrten Wege anzeigen; eher ein Freiheits- als ein Zwangskomplex, da es unbegrenzte Realisationen zuläßt und nur verlangt, daß die funktionellen Bedingungen des Sprachinstrumentes erhalten bleiben“ (aus der deutschsprachigen Zusammenfassung, S. 180).

keit ist prinzipiell gegeben, indem, wie Coseriu sagt, „nicht alles, was normal(sozial) ist, auch funktionell sein muß“.² Im Fall des Deutschen kann dies um so eher der Fall sein, als die neuhochdeutsche Schriftsprache selbst nicht aus einer gesprochenen Sprache, sondern aus geschriebenen Sprachen hervorgegangen ist, so daß das Lautsystem, wie sie es aufweist, eigentlich ein Abklatsch des entstandenen schriftsprachlichen Formensystems war.³

Dieser Gedanke, daß das hochdeutsche Lautsystem seinen Bezug auf mindestens zwei Normensysteme, ein lautlich-funktionelles und ein nichtlautlich-funktionelles, haben kann, bestätigt sich vor allem, wenn man sich mit den einzelnen, heute geltenden Aussprachenormen des Hochdeutschen, die in Aussprachewörterbüchern zu finden sind, beschäftigt. Zwar wird seit *Theodor Siebs*,⁴ oder vielmehr seit *Wilhelm Viëtor*⁵ mit seinem Reformgedanken zum Sprachunterricht, immer wieder beteuert, daß die Aussprache sich niemals nach der Schrift richten könne oder dürfe, sondern nach der gesprochenen Sprache, aber in Wirklichkeit wird konsequent bei allen Ausspracheregeln auf das Schriftbild der einzelnen schriftsprachlich zugelassenen Wortformen zurückgegriffen, so daß z. B. [aɪns] (eins), [ˈhʊnsrʏk] (Hundsrück), [ˈnʊʃəl̩n] (nuscheln), [ˈfʊsəl̩n] (fusseln) als Aussprachenormen vorgeschrieben werden, obwohl in Wirklichkeit nur [aɪnts], [ˈhʊntsʀrʏk], [ˈnʊʒəl̩n], [ˈfʊzəl̩n] gesprochen werden. Diese Tatsache, die sich zwar in diesen Einzelfällen grotesk ausnimmt, dürfte aber gerade die Natur der hochdeutschen Lautnormen recht anschaulich illustrieren. Denn wir wissen z. B. eigentlich noch nicht recht, ob /p/, /t/, /k/ und /b/, /d/, /g/ im Hochdeutschen tatsächlich funktionell bedingte Oppositionen darstellen, obwohl sie in der Phonologie unter Anführung einiger weniger Oppositionspaare wie Bein: Bein, Teich: Deich, Karten: Garten usw. kurzerhand als solche erklärt werden. Jespersen hat bekanntlich in seinem „Lehrbuch der Phonetik“ mit einer vergleichenden Aufstellung zahlreicher Belege darauf hingewiesen, daß es im Hochdeutschen im Gegensatz zum Englischen und Französischen verhältnismäßig wenige Oppositionspaare für /p, t, k/ und /b, d, g/ gibt, worauf sich seiner Ansicht nach das Verwischen dieses Unterschiedes in so vielen Gegenden Deutschlands zurückführen lasse.⁶ Hier sei

² Op. cit. S. 179.

³ Vgl. hierzu Wilhelm Braune: Über die Einigung der deutschen Aussprache (Akademische Rede, Heidelberg 1904), S. 12: „Sie (die hochdeutsche Musteraussprache) hat ihre Grundlage überhaupt nicht im gesprochenen Wort, sondern im geschriebenen: sie ist von Haus aus — wie unsere Schriftsprache überhaupt — eine Sprache nach dem Papier, sie sucht einfach die Wortbilder der historisch gewordenen Orthographie in der gesprochenen Sprache nachzubilden“.

⁴ Theodor Siebs: *Deutsche Bühnenaussprache*, 1. Aufl. (Berlin, Köln, Leipzig 1898), S. 13: „Die Schreibung kann nie und nimmer als Maßstab für die Aussprache dienen.“ — *Wörterbuch der deutschen Aussprache* (Leipzig 1964), S. 12: „Es muß von der Sprechwirklichkeit ausgegangen werden.“

⁵ Quousque Tandem (Wilhelm Viëtor): *Der Sprachunterricht muß umkehren!* (Leipzig 1882).

⁶ Otto Jespersen: *Lehrbuch der Phonetik*, 1. Aufl. (Leipzig 1904), S. 110: „Die brauchbaren Beispiele sind hier (im Deutschen) viel geringer an Zahl als im Französischen und namentlich im

außerdem noch darauf hingewiesen, daß es im Hochdeutschen lexikalisch zugelassene Doppelformen mit p, t, k und b, d, g gibt (knappern-knabbern, Purzelbaum—Bürzelbaum, Borde—Borte, toll—doll, kucken—gucken usw.).⁷ Im krassen Gegensatz zu /p, t, k/ und /b, d, g/ lassen sich dagegen im Hochdeutschen für /r/ und /l/ gleich Hunderte von Oppositionspaaren anführen, und zwar solche, bei denen eine Vernachlässigung der einschlägigen Distinktion wesentliche Kommunikationsstörungen erwarten läßt (Bretter: Blätter, Herrenanzug: hellen Anzug, Rudolf: Ludolf usw.)^{7a}. Hier liegt also offensichtlich wieder ein anderer normativer Bezug, der durchaus funktionell sein dürfte, innerhalb des hochdeutschen Lautsystems vor. Dies scheint auch etwa daraus zu schließen zu sein, daß die Verwechslung der beiden Liquide bei Hörtesten mit Hilfe von sinnlosen Silben im Gegensatz zu den stimmhaften und stimmlosen Verschlusslauten äußerst selten vorkommt, wie es bereits *Hermann Gutzmann* festgestellt hat.⁸

Die phonologischen Einheiten mit ihren distinktiven Merkmalen, die bisher für das hochdeutsche Lautsystem ausgearbeitet worden sind, beruhen nun im wesentlichen auf Analysen derjenigen Normen, die in Aussprachewörterbüchern zu finden sind, die selbst wiederum weitestgehend auf die Schriftnormen ausgerichtet sind. Das heißt, man hat es hier eigentlich weniger mit funktionellen Normen des Hochdeutschen selbst zu tun als vielmehr mit nichtfunktionellen Normen, die „Realisierungen des Systems der geschriebenen Sprache“⁹ wie der ganzen Reihe von Vortragsprachen (öffentliche Reden, Vorlesungen, Predigten, Gerichtsurteile, Nachrichtensendungen, Hörspiele, Theateraufführungen usw.) entnommen sind. Andererseits ist es heute noch nicht klar, welche funktionelle Basis diesen Normen zugrunde liegt, wozu aber eine umfassende Analyse wirklich gesprochener hochdeutscher Sprache notwendig wäre, zu der etwa die Phonometrie bereits seit den dreißiger Jahren Ansätze genommen hat. Das eigentliche phonologische System des Hochdeutschen wird in diesem Sinne erst durch gegenseitige Aufhellung dieser beiden Arten der Normen, die man

Englischen; und es ist auch nicht so sehr wichtig, den Unterschied innezuhalten, weil er in so vielen Gegenden Deutschlands verwischt ist. Obgleich die Witzblätter oft aus solchen Verwechslungen Kapital schlagen . . . , spielt die Unterscheidung zwischen Tenuis und Media faktisch eine sehr geringe Rolle in der Ökonomie der deutschen Sprache.“ In der 2. Auflage (1912) heißt aber der zweite Satz: „und diesem Umstand ist es nach meiner Auffassung von ‚Lautgesetzen‘ (s. Grundfragen Kap. VII, besonders S. 175) wesentlich zuzuschreiben, daß der Unterschied zwischen Tenuis und Media überhaupt in so vielen Gegenden Deutschlands verwischt ist“ (S. 110).

⁷ Vgl. Hermann Paul: *Prinzipien der Sprachgeschichte*, 5. Aufl. (Halle 1920), S. 260: „Wir werden anerkennen müssen, daß eine innere Beziehung zwischen Lautgestalt und Bedeutung nicht vorhanden zu sein braucht, . . .“ (im Zusammenhang mit der Entstehung der Doppelformen).

^{7a} Kennosuke Ezava. Die Opposition stimmhafter und stimmloser Verschlusslaute im Deutschen. Diss. phil. Köln 1969.

⁸ Hermann Gutzmann: Untersuchungen über die Grenzen der sprachlichen Perzeption. In: *Zeitschrift für klinische Medizin* 60 (1906), S. 250.

⁹ Diese Formulierung stammt von Eberhard Zwirner in einem persönlichen Gespräch mit dem Vf.

als „Sprachnormen“ und „Sprechnormen“ bezeichnen könnte, zu finden sein.¹⁰ Dieses Verfahren wird dabei nicht zuletzt zu einer wissenschaftlichen Begründung und Förderung der hochdeutschen Ausspracheregulierung beitragen können.

DISCUSSION

Martens: (nachträglich in schriftlicher Form gemeldet und angenommen; daher keine Antwort des Referenten.)

Es ist völlig richtig, daß die neuhochdeutschen Lautnormen eine gewisse Bindung an die Schriftsprache haben. Aber die Behauptung, das neuhochdeutsche Lautsystem sei ein „Abklatsch des schriftsprachlichen Formensystems“, ist überspitzt. Man kann heute nicht mehr behaupten (wie Braune noch 1904), daß die hochdeutsche Sprache ihre Grundlage überhaupt nicht im gesprochenen Wort habe. (Nach den Arbeiten von Th. Frings und L. E. Schmidt ist das nicht mehr möglich!). Man darf nicht alle Ausgleichbestrebungen für die gesprochene Sprache der Bühne unberücksichtigt lassen (Wanderbühnen, Goethes Regeln für Schauspieler usw.). Andererseits ist es natürlich richtig: die orthographische Form hat einen starken Einfluß auf die Aussprache gehabt und hat ihn auch heute noch, wie übrigens in allen Sprachen — wenigstens halbwegs — phonetisch bedingter Orthographie. Gelegentlich können sich aus solcher Ausrichtung nach der Orthographie eigenartige Differenzierungsnotwendigkeiten ergeben. Nach Duden, Band 6 soll die Zeitschrift von Matthias Claudius „Der Wandsbecker Bote“ mit kurzem offenem, ungespanntem [ε] gesprochen werden — eben weil der Ort zur Zeit Matthias Claudius' mit „ck“ geschrieben wurde. Der Ortsname selbst wird seit 1907 nur noch mit „k“: „Wandsbek“ geschrieben — eben damit er nicht falsch ausgesprochen wird. Hier ist die Entscheidung im Duden, Bd. 6 lediglich nach dem Schriftbild getroffen worden, was seine Berechtigung darin haben mag, daß vielfach eben nach dem Schriftbild gesprochen wird, — auch *gegen* sprachlich richtige Formen!

Die Beispiele „fusseln“ und „nuscheln“, die Herr Ezawa gibt, stimmen in der angegebenen Art nicht: sie werden keineswegs nur mit intervokalisches stimmhaftem, ungespanntem Reibelaut [z] und [ʒ] gesprochen. Diese Ausspracheform gilt in Norddeutschland als Kennzeichen a) für Hineintragen niederdeutscher Artikulationsgewohnheiten ins Hochdeutsche, und b) für sprachlich und meist auch sozial niedrigere Schicht.

Daß die Reihe p t k eindeutig zur Reihe b d g in Opposition steht, wird auch Herr Ezawa nicht bestreiten können. Wir müßten wohl lediglich fragen, ob hier die oft noch angeführten Oppositionsmerkmale „stimmlos — stimmhaft“ weiter gelten können oder ob nicht heute an ihre Stelle die Opposition „gespannt — ungespannt“ getreten ist, besonders im Anlaut.

Natürlich kann man Herrn Ezawa zustimmen, wenn er eine *umfassende* Analyse der gesprochenen hochdeutschen Sprache verlangt; aber man sollte nicht vergessen, daß für das Hallesche „Wörterbuch der deutschen Aussprache“ in dieser Hinsicht schon erhebliche Arbeit geleistet worden ist und daß andererseits die Forschungsvorhaben „Gesprochene Sprache“ und „Hochlautung“ des Instituts für deutsche Sprache (Mannheim) auch einiges erwarten lassen.

Meinhold:

Der von Herrn Ezawa benutzte Normbegriff gestattet es nicht, diachronische Sprach-(Laut)-veränderungen abzubilden. Die drei Hauptphasen im Verlauf solcher Veränderungen (Normabbau

zu Beginn, Realisationsunsicherheit mit rivalisierenden Varianten, neuer Normaufbau) lassen sich allein mit den statistischen Parametern größerer bzw. geringerer Streuung (Dispersion) — neben der Mittelwertbildung — kennzeichnen. Allophonstatistik und Variantenstreuung können angesichts des permanenten phonetischen Variantenbestandes fast aller Sprachen allein Informationen über bestehende oder sich bewegende (abbauende, aufbauende) Normen liefern. Insofern sind Realisationsstatistiken ein unerläßliches Mittel linguistischer Orientierung. — Sodann möchte ich Herrn Ezawa bitten, doch einmal exakt zu definieren, was er unter der erwähnten „funktionellen Basis“ versteht.

Wiede:

Wendet man den Begriff der funktionellen und nichtfunktionellen Normen auf die deutschen stimmhaften und stimmlosen Konsonantenphoneme an, so wie Ezawa das tut, dann muß man zu dem Schluß kommen, daß die Stimmhaftigkeit und Stimmlosigkeit im Deutschen keinen bedeutungsdifferenzierenden Charakter hat. Wodurch unterscheiden sich dann nach Ezawas Ansicht Wortpaare wie Bein—Pein, Garten—Karten, Dorf—Torf, Rauben—Raupen, Egge—Ecke, werden—werten? Die Anzahl solcher Wortpaare ist im Gegensatz zu Ezawas Meinung im Deutschen verhältnismäßig groß.

Vachek:

(Ad Neustupný + Ezawa) Both papers have one feature in common, viz. the stress laid in them on synchronistic vacillation in speech utterances (vacillation pointed out, as early as in 1911, by V. Mathesius). Dr. Neustupný, as a member of the Prague group of today, rightly feels the importance of this fact and wants it to be exactly measured. — The absence of a strict *Sprechnorm*, deplored by Dr. Ezawa, is due exactly to the presence of these vacillating factors. Such presence may be observed, after World War II, also in England, and is a well-known sociolinguistic fact.

Ezawa:

Ad Meinhold: Es sei zuerst auf die von vornherein unzulängliche materielle Basis für Realisationsstatistiken überhaupt hingewiesen, die Herr Meinhold zur Konstatierung des Normenbestandes für unerläßlich hält. Bei diachronischen Untersuchungen ist man bis heute auf schriftliche Abbilder des naiven Sprachbewußtseins angewiesen, die alles andere als direkte Wiedergabe der lautlichen Wirklichkeit (etwa in Tonbandaufnahmen) sind; die funktionelle Bedingtheit des Lautwandels, an die allgemein geglaubt wird, ließe sich unter diesem Gesichtspunkt kritisch überprüfen. Ferner setzen Realisationsstatistiken über eine Sprache mit ihren landschaftlichen, sozialen und individuellen Differenzierungen gewaltige Mengen an systematisch gesammelten Belegen in jeweils repräsentativer materieller Homogenität voraus, ein Grund, weshalb synchronische Variationsuntersuchungen mit eindeutigen statistischen Resultaten nicht ohne weiteres zu erwarten sind (vgl. E. Zwirner und K. Zwirner: *Lesebuch nhd. Texte*, Berlin 1937, Vorwort). — Unter „funktioneller Basis“ verstehe ich die rein funktionellen Bindungen, die die geltenden Lautnormen im Grunde konstatieren, aber nicht diese selbst darstellen (vgl. die Begriffe des „Kommunikativen“ und „Extrakommunikativen“ im Plenarvortrag von G. Ungeheuer auf diesem Kongreß).

Ad Wiede: Die Anzahl der Oppositionspaare für /p, t, k/: /b, d, g/ im Hochdeutschen ist nach eigener Untersuchung tatsächlich verhältnismäßig klein (man kommt im Anlaut nicht über 100, während für /r/: /l/ mindestens 200 Paare anzuführen sind), wobei natürlich außerdem eine Relativierung des gefundenen Formenbestandes vorgenommen werden sollte, indem man die einzelnen Paare auf die Verwechslungsmöglichkeit ihrer Glieder in konkreten syntaktischen

¹⁰ Vgl. S. K. Šaumjan: Die Zweistufentheorie der Phonologie im Licht der modernen Wissenschaftslogik. In: *Phonetica* 16 (1967), S. 135: „Die Notwendigkeit, eine experimentelle Phonologie zu schaffen.“

Zusammenhängen prüft. Trotz dieses weitgehend fehlenden funktionellen Zwangs müssen die bestehenden „Normalformen“ in Paaren mit /p, t, k/ und /b, d, g/ im Hochdeutschen nicht aufgegeben werden, ein Umstand, der zum normalen Status jeder Schriftsprache zu gehören scheint (vgl. hierzu Alfred Schmitt: Die Schallgebärden der Sprache, in: Wörter und Sachen, Bd. XVII, 1936, § 19, 20; § 70, 71).

ad Vachek: Mir ging es darum, auf den wenig beachteten Umstand hinzuweisen, daß die Normen, die dem tatsächlichen Sprechen zugrunde liegen, durch die bisherige Phonologie noch keineswegs erschlossen worden sind und daß die phonologischen Normen, die Phoneme in ihrer Opposition, wie sie etwa vom Hochdeutschen konstatiert werden, weitgehend auf schriftsprachliche Konsequenzen zurückgehen. Die Schwankungen in Realisationen der Sprachnormen (Phoneme) im Sprechen sind nicht gleichzusetzen mit Sprechnormen (bzw. Lautungsnormen), deren eindeutige positive und negative funktionelle Ordnung es heute zu erforschen gilt (vgl. Plenarvortrag von G. Ungeheuer: *Kommunikative und extrakommunikative Betrachtungen in der Phonetik*).

ZUR PERZEPTION UND PHONOLOGISCHEN INTERPRETATION DES *O NACH LABIALEN UND VELAREN IM NIEDERSORBISCHEN

HELMUT FASSKE*

In den niedersorbischen Dialekten ist die Qualität des etymologischen *o von dessen Stellung im Wort abhängig. In akzentuierter Silbe wird jedes etymologische o nach Labialen und Velaren qualitativ modifiziert. Es hat auch an Versuchen nicht gefehlt, den Lautwert der einzelnen Kontinuanten von etymologischem o nach den Labialen und Velaren akustisch-artikulatorisch zu bestimmen und phonologisch zu interpretieren. Einige Forscher glauben, an Stelle des etymologischen o nach Velaren und Labialen ein von allen anderen Vokalen deutlich unterscheidbares labialisiertes \dot{y} bzw. in anderen Dialekten ein labialisiertes \dot{e} zu hören. Allerdings leugnen auch sie nicht die Tatsache, daß etymologisches o nach Labialen und Velaren oft durch nichtlabialisiertes y bzw. e vertreten wird (*gylc* „der Junge“ — *dwer* „der Hof“). Die phonetische Wertung der Kontinuanten von etymologischem o nach Labialen und Velaren in den niedersorbischen Dialekten und die davon abhängige phonologische Interpretation, die sich im weitestgehendem Maße oder ausschließlich auf das Gehör des Explorators stützt, ist nicht frei von subjektiver Auffassung und Entscheidung des Einzelnen. Es ist für einen Sprachwissenschaftler, der nicht von Kindheit an und aus eigener sprachlicher Erfahrung die untersuchte Sprachvariante kennt, äußerst schwierig, wenn nicht gar unmöglich, sein eigenes Lautsystem als Filter des Aufgenommenen auszuschalten und alle, oftmals ad hoc formulierte Postulate zu eliminieren.

Eine objektiv gesicherte Entscheidung über den phonologischen Wert der genannten Kontinuanten in den niedersorbischen Dialekten muß von der Existenz bzw. Nichtexistenz einer akustisch deutlich wahrnehmbaren Opposition nichtlabialisiertes y — labialisiertes \dot{y} bzw. nichtlabialisiertes e — labialisiertes \dot{e} und einer damit verbundenen sinnunterscheidenden Funktion dieser Opposition in gleichen Positionen ausgehen. Da das Gehör des Sprachwissenschaftlers für eine objektive Entscheidung nicht ausreicht (und experimentell-phonetische Mittel stehen dem Forscher im Feld nicht zur Verfügung), so habe ich einen Versuch vorbereitet und durchgeführt, der weitgehend eine subjektiv bestimmte Entscheidung ausschließen sollte.

Ich habe zunächst ein Wortpaar zu einer gewürfelten Reihe zusammengestellt,

* DAW zu Berlin, Institut für sorbische Volksforschung Bautzen.

das sich lediglich etymologisch durch den Stammvokal unterscheidet. Die erste Wortreihe enthielt die Wörter *wen* „hinaus“ (mit etymologischem *e* nach einem Labial) und *wen* „er“ (mit etymologischem *o* nach einem Labial).

Diese Reihe habe ich in einem Ort im Dialektgebiet, in welchem jedes etymologische *o* nach Labialen und Velaren — nach Meinung einiger Forscher — als labialisiertes \hat{e} gesprochen wird, durch einen ortsansässigen Dialektsprecher in heimischer niedersorbischer Mundart auf Tonband sprechen lassen, und zwar so, daß auf dem Tonband nur eine Reihe gleichlautender, aber etymologisch und semantisch unterschiedlicher Wörter erschien, nämlich:

wen (er) — *wen* (er) — *wen* (hinaus) — *wen* (er) — *wen* (hinaus) . . . Um festzustellen, ob die Sprecher dieses Dialekts selbst die beiden Bedeutungen der Wörter in dieser gewürfelten Folge nach Gehör unterscheiden können — was bei positivem Erfolg die Existenz der Opposition *e* (nichtlabialisiert) — \hat{e} (labialisiert) und damit die sinnunterscheidende Funktion dieser Opposition bewiesen hätte — habe ich anderen, bei der Aufnahme nicht anwesenden ortsansässigen Gewährspersonen die auf Tonband aufgenommene Reihe vorgespielt mit der Bitte, die Bedeutung der gehörten Wörter zu nennen.

Die erste Gewährsperson, der ich die Bedeutung der Wörter in dieser Reihe vorher nicht genannt habe, gibt das Gehörte ausnahmslos mit „hinaus“ wieder. Da aber die Reihe 7mal das Wort mit der Bedeutung „er“ enthält, so ergeben sich 7 falsche Entscheidungen. Auch die zweite Gewährsperson gibt bei gleichen Bedingungen die Bedeutung der Wörter dieser Reihe ausnahmslos mit „hinaus“ wieder.

Der dritten Versuchsperson erklärte ich vorher, daß auf dem Tonband zwei Wörter aufgenommen sind, eines in der Bedeutung „hinaus“, das andere in der Bedeutung „er“. Sie sollte in jedem einzelnen Fall entscheiden, ob das gehörte Wort die Bedeutung „hinaus“ oder die Bedeutung „er“ habe. Das Resultat war negativ. Die Versuchsperson traf 7 Fehlentscheidungen. 5mal wurde die Bedeutung „hinaus“ angegeben, wo „er“ erwartet werden mußte, 2mal wurde fälschlicherweise „er“ angegeben, wo die Wiedergabe „hinaus“ lauten mußte.

Schließlich habe ich zum Abschluß meine Gewährsperson, die mir die Wortreihe auf Tonband gesprochen hatte, gebeten, selbst zu bestimmen, wann sie „hinaus“, wann „er“ gesagt habe. Auch dieser Versuch war negativ. Es wurden 8 falsche Entscheidungen getroffen. Das Ergebnis war also eindeutig: Die Dialektsprecher unterscheiden nicht zwischen etymologischem *e* und \hat{e} , das aus etymologischem *o* entstanden ist.

Ich unternahm noch einen zweiten parallelen Versuch im Dialektgebiet, in dem jedes etymologische *o* nach Labialen und Velaren nach Meinung einiger Forscher zu labialisiertem \hat{y} geworden sei. Hier nahm ich eine Wortreihe auf Tonband auf, die aus den beiden Wörtern *myta* „die Löhne“ (mit etymologischem *y*) und *myta* < *mota* „er weiß“ (mit etymologischem *o*) bestand. Auch in diesem Falle fiel das Ergebnis absolut negativ aus. Die Dialektsprecher selbst können akustisch nicht

zwischen etymologischem *y* und \hat{y} , das aus etymologischem *o* entstanden ist, unterscheiden.

Es bliebe nur noch die Frage offen, ob das geübte Ohr des geschulten Dialektologen zwischen diesen Konsonanten zu unterscheiden vermag. Ich habe einen meiner Kollegen gebeten, in präziser phonetischer Umschrift die aufgenommenen Wortreihen wiederzugeben. In der Reihe *wen* wurde der Vokal von 13 möglichen Fällen 10mal als nichtlabialisiertes ϵ wiedergegeben, 3mal nur als labialisiertes $\hat{\epsilon}$, davon jedoch zweimal für etymologisches *e*, einmal nur für etym. *o*. In der Reihe *myta* ist der Vokal kein einziges mal als labialisierter Laut der mittleren Reihe wiedergegeben worden.

Die Annahme einer Opposition nichtlabialisierter Vokal — labialisierter Vokal, in der sich die Unterscheidung von etymologischem *e* bzw. *y* und etymologischem *o* nach Labialen und Velaren widerspiegeln, beruht auf einer Täuschung, die durch die Kenntnis der Etymologie und Sprachgeschichte verursacht wurde. Ein Versuch, der subjektive Auffassungen weitgehend ausschloß und kontextliche oder situationsgebundene Unterscheidungsmöglichkeiten eliminiert, dürfte, so glaube ich, bewiesen haben, daß auch die im Feld betriebene und notgedrungenermaßen auf technische Hilfsmittel verzichtende Phonetik über geeignete objektive Methoden verfügt, um sprachliche Laute in notwendigem Maße zu identifizieren und ins phonologische Lautsystem verläßlich einzuordnen.

DISCUSSION

Hamm:

(ad Faske + Kelly)

1. Calling attention to similar works by Matejka + Magrer on Serbo-Croat forms.
2. Is there any morphophonemic connection between ${}^{\circ}CC \dots$ and CVC^* (${}^{\circ}$ and * being signs for palatalisations)?

LA DESCRIPTION PHONOLOGIQUE DES SYSTÈMES PROSODIQUES*

GEORGES FAURE**

En dépit du fait que certains linguistes considèrent encore que la prosodie ne relève pas de la description phonologique, nous sommes de plus en plus convaincu — avec de nombreux chercheurs — que la structuration prosodique d'un énoncé relève d'un système tout aussi rigoureux et tout aussi économique que le système phonématique.

Nos recherches récentes et les résultats — encore partiels et inédits — de divers travaux dont nous assurons la direction, nous ont, en effet, conduit à penser qu'il existe pour chaque langue un système prosodématique spécifique, définissable en termes d'unités discrètes, en nombre déterminé et constituées, à tous les niveaux d'analyses, par des faisceaux de traits distinctifs réalisés et perçus simultanément. Ces unités se situent, comme les phonèmes, dans les limites d'une marge de tolérance définie par des seuils, qu'il s'agisse de leur réalisation par le locuteur, ou de leur perception et de leur interprétation fonctionnelle par l'auditeur; la commutation de ces unités, sur un point déterminé de la chaîne parlée ayant pour effet de changer l'identité du signifié psychologique.

C'est ce que nous essaierons de démontrer et d'illustrer à l'aide d'un certain nombre d'exemples précis empruntés à la langue française.

Nous montrerons, d'autre part qu'en dépit de certains traits originaux que nous nous efforcerons de dégager, le comportement fonctionnel des systèmes prosodiques est, dans une large mesure, comparable à celui des systèmes phonématiques.

Il en résulte que l'on peut parler, selon nous, en toute sérénité de prosodèmes (voire, pour une analyse plus serrée, de dynèmes, de chronèmes et surtout de tonèmes) dont l'inventaire peut être établi pour une langue donnée, selon les mêmes critères que pour les phonèmes.

Nous montrerons, d'autre part, que l'étude du rendement fonctionnel des oppositions prosodématiques peut se faire à l'aide de techniques comparables à celles

* Le texte complet sera publié dans la revue *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 23, 1970, H. 3.

** Chaire de Phonétique de la Faculté des Lettres et Sciences Humaines de l'Université d'Aix-Marseille.

qui permettent de préciser le rendement fonctionnel des oppositions phonématiques.

Nous soulignerons enfin que l'on peut, comme pour les phonèmes, étudier les effets de ce rendement sur le degré de stabilité ou d'instabilité des structures phonologiques d'une langue donnée.

On peut également étudier, dans cette perspective, les effets des conditions de réalisation et des conditions de perception des unités prosodématiques. (Une opposition se maintenant d'autant plus solidement à rendement fonctionnel comparable, que cette opposition est plus facile à réaliser ou à percevoir). Ce dernier fait semble exclure, entre autres, l'idée d'une modification du signifié en fonction de chaque variation si infime soit-elle, de la réalisation prosodique affectant telle ou telle unité de l'énoncé.

SOME ASPECTS OF HIGH-ACCURACY ANALOG FUNDAMENTAL - FREQUENCY RECORDING

MIROSLAV FILIP*

Fundamental frequency recording represents an area in which the interests of phonetics join with those of musical acoustics. However, a quantitative difference between these two aspects still remains being affected, in general, by the well-known principle of uncertainty (1, 2): while considerably higher *time* resolution, of the order of few milliseconds, is mostly required in speech research, in musicology, on the other hand, an extreme *frequency* resolution is primarily claimed.

It is hoped that some comment concerning high-accuracy frequency recording, given by a musicologist, may perhaps be of interest to phoneticians.

In speech intonation research, as well as, for example, in ethnomusicology, a durable analog record of fundamental frequency as a function of time proved to be a most valuable aid. Disregarding for the present purpose the spectral, autocorrelation, and digital methods we have at disposal two classes of simple analog means: first, integration methods which, however, are inferior in recording rapid changes of instantaneous frequency; second, timing methods, first described by Grützmacher and Lottermoser (3), which make use of particular timing waveforms to determine the duration of each fundamental period.

It is the logarithmic frequency scale that has been found most advantageous in several respects; one additional will be given later. Apparently, in the field of our interest, only the methods yielding the logarithm of the *instantaneous* value of the input quantity are to be taken into account.

I shall now briefly discuss three basic methods for obtaining the logarithmic display at the output of a frequency recorder.

- (a) application of a logarithmic four-terminal network with inherent logarithmic characteristic (4, 5)
- (b) approximation by a piecewise linear (polygonal) function (6, 7)
- (c) approximation by a linear combination of exponential functions (8, 9)

(a) Logarithmic networks are based on more or less accurate logarithmic relation between two physical quantities in vacuum tubes or semiconductors. The former can

* Department of Musicology, Univerzita Komenského, Bratislava.

hardly be exploited economically in the problem under consideration, due to large drift. The latter are easier to use, despite of the undesired temperature dependence. There were many semiconductor logarithmic networks described in the literature, covering often as much as seven decades of the input quantity with an accuracy of several per cent. For the present purpose one decade would be far enough but a higher accuracy is required. I would like to present a very simple circuit giving good logarithmic response over two octaves with unselected general-purpose germanium diodes. The average error of approximation was estimated to be about one half per cent.

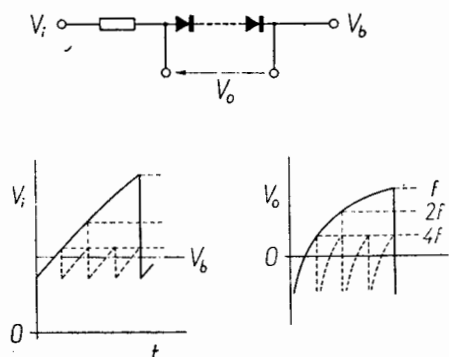


Fig. 1. Simple quasilogarithmic network transforms the initial, approximately linear, portion of an exponential function into logarithmic function within the range from f to $4f$. V_i — input voltage from an instantaneous frequency meter based on (3), V_o — logarithmic output fed to a differential amplifier with high input impedance, V_b — fixed bias voltage. Instead of one, several diodes are connected in series to increase V_o .

(b) Piecewise linear approximation based on nonlinear voltage dividers known from diode function generators is to be preferred when large frequency range (up to two decades) is needed along with relatively suppressed claims to accuracy. In high-precision instruments this technique tends to become somewhat complicated. It is often employed in biological research, for example, in cardiometers and in pulse rate meters for electromyography, usually approximating the hyperbolic function (10, 11).

(c) Third fruitful approach to the problem is in approximating the logarithmic function by a linear combination of exponential functions with different decay constants which are available with high accuracy and reproducibility by charging capacitors in simple RC circuits connected to a common output. Very good results were obtained in wide frequency range instruments.

On the other hand, it may be shown that in two-octave range the improvement resulting from application of *two* different time constants, as compared to a single RC circuit, is not substantial. Approximation by a *single* exponential curve (3, 12, 13) is, of course, only very rough for ranges greater than, say, sixteen semitones. A phonetician and a musicologist would naturally welcome very large frequency ranges

operated without switching, even more than two octaves. However, the errors in measurement due to the limited accuracy of analog display will generally be expressed as a constant fraction of the full scale deflection. The experimenter is thus forced to reduce reasonably the display range. A compromise must be made in order to improve the accuracy without sacrificing a sufficiently wide range. (The instruments are almost always designed as multirange ones.)

Fortunately, as anticipated above, the logarithmic frequency scale has an extra advantage with regard to the problem just discussed. I would like to suggest that *one-octave* range could be entirely satisfying since melodies exceeding that interval may simply be recorded by repetition employing successively all contiguous one-octave bands occupied by the melody as a whole. Finished records are then properly combined (fastened, glued) in parallel with respect to the time axis so that they form a composite record having just the range needed by the melody but, of course, with the accuracy valid for one-octave range.

Then, *one single* exponential function will be sufficient for the high-accuracy approximation. The calculated nonlinearity of the semitone scale does not exceed 5 cents in terms of musical intervals or, in other words, maximum error amounts approximately 0.4% of the full scale deflection, assuming that the time constant equalled the reciprocal of the center frequency of the range.

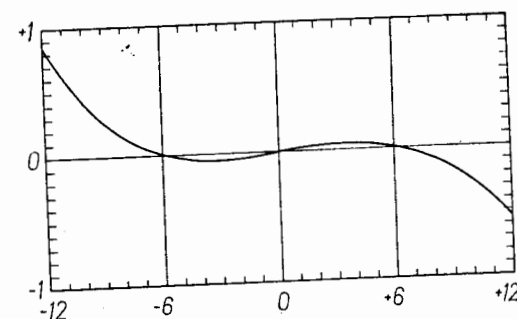


Fig. 2. Departure from linearity of the semitone scale obtained by an exponential approximation with the time constant equal to the period of the center frequency. Horizontal axis: log frequency in semitones re center frequency; vertical axis: deviation of the actual scale from a linear one, in semitones. This may be taken into account when calibrating the instrument or, in one-octave range, it may be neglected being less than ± 5 C.

REFERENCES

1. Meyer—Eppler W., „Tonhöhenschreiber“, *Z. Phonetik*, 2 (1948), pp. 16—38.
2. Filip M., „Frekvenčné merania a tónová sústava“, (Frequency Measurements and the Tone System), in: *Nové cesty hudby II* (New Ways of Music II), Praha 1969 (in print).
3. Grützmacher M., and Lottermoser W., „Über ein Verfahren zur trageheitsfreien Aufzeichnung von Melodiekurven“, *Akust. Z.*, 2 (1937), Sept., pp. 242—248.
4. Obata J., and Kobayashi R., „A Direct-Reading Pitch Recorder and Its Applications to Music and Speech“, *J. Acoust. Soc. Am.*, 9 (1937), Oct., pp. 156—161.
5. Carré R., Lancia R., Paillé J., and Gsell R., „Étude et réalisation d'un détecteur de mélodie pour analyse de la parole“, *L'onde électrique*, 43 (1963), No. 434, pp. 556—562.
6. Mansford H. L., and Khan K. M. I., „An Amplitude/Frequency Response Display Using a Ratio Method“, *Electron. Engng.*, 30 (1958), No. 367, pp. 541—544 (Part I).
7. Nathan A., „Linear and Nonlinear Interpolators“, *IEEE Trans. Electron. Computers*, EC-12 (1963), No. 5, pp. 526—532.
8. Fauerskov S. E., „The Frequency Response Tracer Type 4709“, *Brüel and Kjaer Technical Review* 1965, No. 3, pp. 20—33.
9. Alcock R. N., „A Logarithmic Time Base“, *Electron. Engng.*, 36 (1964), No. 436, pp. 404—406.
10. Andrew A. M., and Roberts T. D. M., „A Pulse-Interval Meter for Measuring Pulse Repetition Frequency“, *Electron. Engng.*, 26 (1954), No. 321, pp. 469—474, No. 322, pp. 543—547.
11. Kay R. H., „A Reciprocal Time-Interval Display Using Transistor Circuits“, *Electron. Engng.*, 37 (1965), No. 450, pp. 543—545.
12. Kallenbach W., „Eine Weiterentwicklung des Tonhöhenschreibers mit Anwendungen bei phonetischen Untersuchungen“, *Akust. Beih.*, H. 1 (1951), pp. 37—42.
13. Tove P. A., and Czekajewski J., „Pulse Period Meter with Short Response Time. Applied to Cardiotachometry“, *Electron. Engng.*, 36 (1964), No. 435, pp. 290—295.

"TONEME 3" IN NORWEGIAN

KNUT FINTOFT

The Norwegian toneme system is usually regarded as being binary and toneme oppositions only occur in syllables with primary stress, followed by one or more unstressed syllables, or when a stressed monosyllable is followed by an unstressed word. Monosyllables have traditionally been regarded as having toneme 1.

Vanvik^{4,5} has raised the question, if there exists a third toneme in Norwegian. According to him the traditional opinion that monosyllables have toneme 1, seems to be based chiefly upon historical considerations. He points out that a stressed monosyllable followed by one or more unstressed syllables can contrast with polysyllabic words (with toneme 1 or 2), and that the primary acoustical difference between toneme 1 and "toneme 3" is that of $F_0^{6,7}$. Mo³ also sets up a third toneme for stressed monosyllables. Haugen² does not agree with Vanvik's interpretation, and Borgström¹ has pointed out that the third toneme may be regarded as an allotone to toneme 1.

A listening test and acoustical analysis have been carried out. The results from three town dialects will be given here. It is assumed that in disyllabic words the tonemes may be recognized in the stressed syllable. A monosyllabic word may therefore be compared with the stressed syllable in toneme 1 and 2 words. The name *Liv* was read twice in the frame *navnet er Liv* by three subjects from each of the towns of Oslo, Ålesund, and Trondheim, and made the basis for a listening test. The words were presented in random order to listener groups (167 subjects) speaking the same dialects. The listeners were about 16 years old, not phonetically trained and had never taken part in similar experiments. They were told that they would hear the first part of the words '*livet* or '*live* spoken in different dialects, and were asked to mark on forms to fill in which of the two words they thought the segments were cut from. They had no idea that the original word was simply *Liv*.

The responses were largely dependent upon both the speaker and the listener group. When speaker and listener used the same dialect the decision for toneme 1 was very significant for Oslo and Ålesund, whereas no significance for either of the tonemes could be detected for Trondheim (see table). In fact, none of the Trondheim speakers were significantly decided to use either toneme 1 or toneme 2 by any of the listener groups. Similarly, the combination Oslo speaker and Trondheim listeners showed no significant response for toneme 1 or toneme 2. For comparison, all the

listener groups significantly identified the tonemes in the words 'livet and 'live spoken by the same subjects, even when they did not hear the final vowel. Fig. 1 shows the F_0 tracings for the words *Liv*, 'livet, and 'live spoken in the three dialects.

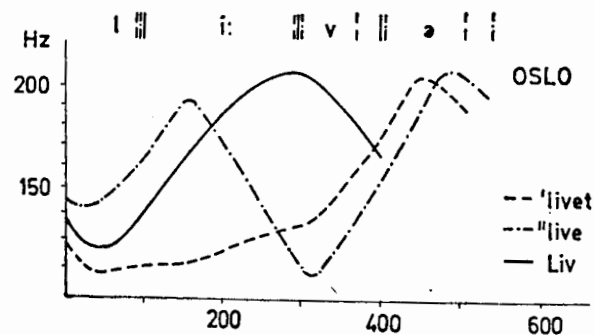


Fig. 1.

The duration and the frequency pattern are average values for each sound. At the first rate the curves for the monosyllables deviate rather much from toneme 1 in Oslo and Trondheim dialects, and they seem to be more like toneme 2. But the most important cue for toneme 2 in these dialects seems to be the falling F_0 from the be-

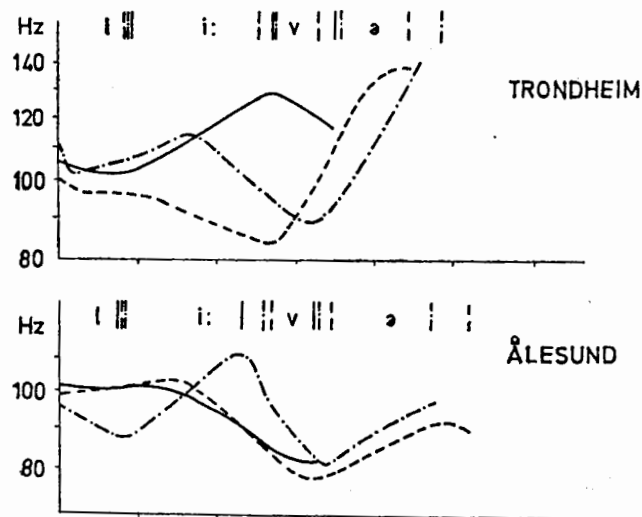


Fig. 2.

ginning of the stressed vowel which does not occur in the monosyllable. In Oslo dialect, however, the raising F_0 during the entire stressed vowel corresponds to toneme 1. In Trondheim, on the other hand, the F_0 pattern for the monosyllable does not correspond to either of the two tonemes.

From the perceptual point of view the results may indicate that there may be a need for "toneme 3" in Trondheim dialect. Only in Ålesund dialect do Trondheim listeners identify monosyllables as having toneme 1. In this case the F_0 -curve for *Liv* and 'livet are practically identical. It is not time to discuss the interaction between the dialects any further, as this problem is much more complicated.

From about 3000 decisions (all the material in bulk) there is significance on the 0.1% level for the monosyllables to be identified as having toneme 1. If the monosyllables had "toneme 3", and "toneme 3" deviates from both the other tonemes, none of the two would be preferred by the listeners. That means that about 50% of the responses would be given to each of the tonemes. The decisions are, as seen, largely dependent both on the individual speaker and the speaker's and listener's dialect, but by and large the monosyllables are identified as having toneme 1, and "toneme 3" may be regarded as an allotone to toneme 1.

A monosyllable, *Liv*' identified as having toneme 1 or 2

		Listeners					
		Oslo		Ålesund		Trondheim	
		1	2	1	2	1	2
Readers	Oslo	196	140	202	122	178	164
	Ålesund	190	146	203	121	218	124
	Trondheim	180	156	168	156	165	177

Level of significance for toneme 1

		Listeners		
		Oslo	Ålesund	Trondheim
Readers	Oslo	1 %	0,1 %	—
	Ålesund	5 %	0,1 %	0,1 %
	Trondheim	—	—	—

REFERENCES

- Borgström, C. H.: *Studia linguistica* 16, 34—37 (1962).
- Haugen, E.: *Monatshefte* 55, 157—161 (1963).
- Mo, E.: Forhand. Vidsk. selsk. skr., Oslo 1922.
- Vanvik, A.: *Maal og minne* 92—102 (1956).
- Vanvik, A.: *Studia linguistica* 15, 23—28 (1961).
- Vanvik, A.: *Studia linguistica* 17, 47—53 (1963).
- Vanvik, A.: *Phonetica* 10, 165—173 (1963).

DISCUSSION

Vanvik:

The term "toneme 3" should not be used at all, since a monosyllable cannot possibly form a minimal pair with a polysyllable by means of tonemic opposition.—Very much depends on the pronunciation of the test word. Since it carries an intonational nucleus of a falling type, the results of the investigation say nothing about the majority of monosyllables, viz. those occurring in non-nuclear position and those carrying an intonational nucleus of a rising type.

NOTE ON THE PROSODIC FEATURES AS MEANS OF FUNCTIONAL SENTENCE PERSPECTIVE

JAN FIRBAS*

In the theory of functional sentence perspective (= FSP),¹ a vexed question is presented by the concept of transition. Somewhat modifying I. P. Raspopov's view, I have found (in an article published in *Philologica Pragensia* 8/1965) that in their non-marked use on the level of FSP, the temporal and modal exponents of the finite verb (TME's) start building up, on a foundation provided by the thematic elements, the very information to be conveyed by the sentence or clause. By themselves, or in co-operation with other elements, they mediate between the theme (constituted by elements carrying the lowest degrees of communicative dynamism [=CD]) and the rheme (constituted by elements carrying the highest degrees of CD): they carry the lowest degree of CD within the non-thematic section of the sentence or clause.² In their marked use, they function either in theme proper or rheme proper, signalling a very special kind of contextual dependence. The question may be raised to what extent an inquiry into the prosodic features of the finite verb (= f.v.) would corroborate this view, which points to a remarkably high coincidence between FSP on the one hand, and the semantic and grammatical levels on the other.

There cannot be a perfect correspondence between the gamut of prosodic weight (unstressed — partially stressed — stressed — bearing a nucleus [cf. A. C. Gimson's *Introduction* . . . p. 244]) and the gamut of CD (ranging from theme proper over transition to rheme proper). Such a perfect correspondence would not but impede language in fully coping with all its tasks. It is well-known, e.g., that the theme may undergo prosodic intensification. The same, in fact, applies to transition. Such intensification, however, must remain within certain limits. The respective elements would cease to be thematic or transitional if not followed by a functionally weightier prosodic feature. (A low rise occurring after a fall within one and the same distributional field of CD would not constitute such a feature. This is perhaps the chief qualification we have to bear in mind when using the above-mentioned prosodic gamut.)

* Department of English, University of Brno.

¹ For an outline of my conception of the theory of FSP, see my paper delivered at the Xth International Congress of Linguists in Bucharest 1967.

² A misleading omission has occurred in the summary of my paper. The beginning of line 5 of par. 2 should read: within the non-thematic section of the sentence.

I will now offer the main results of an inquiry into the prosodic features of 419 f.v.'s examined in the light of FSP. The material has been drawn from P. A. D. MacCarthy's *English Conversation Reader* (1956), pp. 24—45. The inquiry is to be published in full in *Brno Studies in English* 7.

It has revealed that there are at least six types of communicative function the English f.v. can perform. Illustrative examples follow: (i) 'What did you say?—25, ex. 16; (ii) 'They were booked up' too, really, ... —32.28; (iii) The proprietor was 'most friendly — ... 33.05; (iv) Then I retired to a 'seat in the ,park and spent 'half an ,hour or so ... —32.06; (v) ,Well, I'll ,tell you. —32.14; (vi) 'Well that ,does sound nice. — 33.08.

T(type) (i) is represented by 1 case (0,2%) in my material. The f.v. is entirely thematic. *T*(ii). 13 cases (3,1 %). The f.v. is thematic with the exception of the TME's, which constitute transition proper. Both *T*(i) and *T*(ii) are excluded from the sphere of stress bearers and nucleus bearers. *T*(iii). 122 cases (29,1 %). The f.v. is non-thematic, functioning entirely within transition proper. With the exception of 3, all cases are excluded from the sphere of stress bearers and nucleus bearers. *T*(iv). 212 cases (50,6 %). The f.v. is non-thematic, functioning in transition proper only through its TME's, but not expressing rheme proper. In 100 cases it functions as a stress bearer, in 13 cases as a nucleus bearer. *T*(v). 61 cases (14,6 %). The f.v. is non-thematic; it expresses rheme proper, though through its TME's it simultaneously functions in transition proper. 56 cases out of the 61 (91,8%) are nucleus bearers. (Though not nucleus bearers, the remaining cases appear as the prosodically weightiest elements within the respective sentence or clause structures.) *T*(vi). 10 cases (2,4 %). The f.v. expresses rheme proper, this time through its TME's; otherwise it is thematic.

Borderline cases can naturally be established. (They are to be discussed in *BSE* 7.) They do not, however, invalidate the following conclusions. The f.v. displays an unmistakable tendency to function as a transitional element. In regard to degrees of CD, it is, in an overwhelming majority of cases heterogeneous. On the prosodic level this heterogeneity becomes especially evident when the TME's are separate words.

Now there are 127 cases in which the TME's are expressed by separate words (auxiliaries). In 10 of these cases they function as rheme proper and are nucleus bearers, in 1 case as theme proper and are unstressed; in 117 cases they function as transition proper and occur outside the nucleus sphere. If to the last mentioned group, the 114 cases are added in which the f.v. form has no auxiliary, but occurs within transition proper, the number of cases in which transition proper is expressed by at least one separate word is raised to 241, i.e. 58,7 per cent of the total number of all the examined cases. The total number of nucleus bearers among these 241 cases amounts to 4.

All this is not at variance with the assignment to transition proper of those TME's that are expressed by bound morphemes. The weight of the prosodic feature of a word consisting of more than one morpheme naturally depends on that morpheme which carries the highest degree of CD within the examined word.

To sum up. It is hoped that the offered inquiry bears out the conclusion that the TME's show a remarkably high degree of coincidence between FSP on the one hand and the semantic and grammatical sentence structures on the other. It is also hoped that it testifies to the usefulness of the concept of transition proper in functional sentence analysis.

Table 1

Type	The Finite Verb of the Sentence or Clause										Total Number of Cases	
	The Sphere of Function in FSP of ¹		Prosodic Characteristics of									The Non-Thematic Section of the Sentence or Clause ² (not including the finite verb) Number of non-verbal nucleus bearers occurring within the non-thematic section that are functionally weightier than the finite verb
	the TME's of the finite verb form	the finite verb form to the exclusion of its TME's	the auxiliaries (if present) of the finite verb form				the finite verb form to the exclusion of possible auxiliaries					
			Number of Cases									
		unstressed	partially stressed	stressed	bearing a nucleus	unstressed	partially stressed	stressed	bearing a nucleus			
I	theme proper	theme proper	1	—	—	—	—	1	—	—	one	1
II	transition proper	theme	2	—	—	—	8	5	—	—	at least one	13
III	transition proper	transition proper	8	—	—	—	117	2	2	1	at least one	122
IV	transition proper	transition	78	1	8	1	77	22	100	13	at least one	212
V	transition proper	rheme proper	17	—	1	—	3	2	—	56	none	61
VI	rheme proper	theme ³	—	—	—	10	3	—	—	1	none	10

¹ Strictly speaking, this column records how the finite verb function in FSP through its semantic content. ² The term 'non-thematic' is always to be understood here in regard to the distributional field in which the examined verb functions. ³ Provided this component occurs. Cf. the higher number in the column giving the prosodic characteristics of the auxiliaries.

Table 2

Type	Characteristics of the Finite Verb (Cf. Table 1)		Total	Percentage		
I	theme proper	theme proper	1	0,2	0,2	0,2
II	transition proper	theme	13	3,1	3,1	} 97,4
III	transition proper	transition proper	122	29,1	} 79,7	
IV	transition proper	transition	212	50,6		
V	transition proper	rheme proper	61	14,6	14,6	
VI	theme proper	theme	10	2,4	2,4	
			419	100,0	100,0	100,0

DISCUSSION

Storms:

In a number of instances a special intonation is linked with the introduction of a special grammatical word (He is *most* friendly; it *does* sound nice), whereas in other cases there is no special grammatical word. Now my question is: Can you say anything about the percentages of the introduction or absence of grammatical means?

Wode:

It seems to me that statistical findings should be evaluated in relation to innate stress placement. For instance in

I can *only* see *yóu*

only refers to *you*. In

I can *only* *sée* you

only refers to *see*. This should prove important for FSP.

Firbas:

ad Storms: The laws of the interplay of means of FSP determine under what conditions certain grammatical elements carry certain degrees of CD. As the location of the nuclei is determined by these laws, I hold that even the emotive intensification of prosodic features, which may be considerable, must not interfere with the function the sentence performs in FSP.

The last intervention concerns the relation between grammatical structure, FSP and the prosodic means expressing emotiveness.

ad Wode: It would be worth finding out to what extent Dr. Wode's term 'inherence' could be made to correspond to my term 'coincidence'. I have not collected enough examples of his type to be in a position to give a satisfactory reply. But as to *only*, I regard it as a rhematizer, i.e. as an element which signals the element to which it refers as rhematic.

Firbas:

In my previous researches into FSP, carried out almost exclusively in the non-prosodic sphere, I have by no means confined myself to the finite verb. In the present paper, I have concentrated my attention on the TME's, for as they present a point of remarkably high coincidence between FSP on the one hand, and the semantic and grammatical sentence structures on the other, they appear to be a good starting point for an inquiry into the means of FSP in the prosodic sphere.

INTRA-SYLLABIC PITCH MOVEMENT IN COMMUNICATIVE AUSTRALIAN ENGLISH UTTERANCE MORPHEMES

E. H. FLINT

One of the fundamental characteristics of M. A. K. Halliday's system of describing British English (RP) intonation is a recognition of pitch movement as a feature of intonation.¹ All of the five tones of his primary system, and most of those of his secondary systems, exhibit pitch movement of some kind. Halliday moreover relates the description of intonation to that of rhythm; but rhythm is for him a complex conception, resting on a distinction between 'strong' and 'weak' syllables, where strength and weakness depend not merely on intensity. What in British English are usually called "four degrees of stress" are rather "structurally identified syllable classes whose exponents are marked by contrast not only (if indeed at all) in intensity but also in pitch and duration." Halliday also recognizes that intonation may be grammatically (and therefore meaningfully) relevant.

Early descriptions of American English intonation are in general based upon the recognition of relative contrastive pitch levels, termed 'tonemes', and terminal contours. The description of intonation is related to other phonological characteristics — rhythm, pause, length and tempo of articulation.

The recognition of relative contrastive pitch levels is useful to the linguist for describing intonation. Too great emphasis upon it may tend to obscure the important phenomenon of pitch movement. This is evident in the hypothesis of Trager and Smith that "pitch as used in language is heard around a limited number of points rather than as a continuum".² Trager does indeed, after recognizing four pitch phonemes, admit the existence of "variations within" the four levels of pitch, which he terms allophones and describes by means of diacritic symbols.

How can instrumental analysis elucidate the problems of the auditory linguistic analysis of intonation, in relation to loudness and length? How is the auditory perception of these complex features related to the communication of meaning?

The investigation of these problems for the purposes of this paper was confined to the study of the prosodic features of phonemic syllables constituting monosyllabic

¹ M. A. K. Halliday, "The Tones of English," *Archivum Linguisticum*, 15, 1 (1963), 1—28.

² G. L. Trager and H. L. Smith, *An Outline of English Structure* (Washington: American Council of Learned Societies, 1957. Original edition, 1951), p. 41.

utterance and sentence level morphemes in continuous communicative utterances spoken at normal conversational tempo (e.g. affirmatives such as *yes*, introducers such as *oh*, and intensifiers such as *pretty*), and to certain simple phrasal combinations in which these occurred (e.g. *pretty good*).³ These items were particularly suitable for the purpose, because they were more easily segmentable from the continuous communicative utterances in which they were studied, and because they are characteristically intonation bearers.

After a rigorous grammatical and lexical analysis of all the utterances, an auditory analysis of the pitch, loudness and length characteristics of the particular items in their utterance phonological setting was made. Finally, an acoustic frequency-intensity-duration analysis was made, and the results of the auditory and acoustic analyses were compared.

The apparatus used was an IDFF (Intensity-Duration-Frequency of Fundamental) Analyser. This presents a simultaneous display of the acoustic characteristics of an utterance up to 2.5 seconds in length, measurable by scales, and precisely identifiable with respect to the successive phonetic segments to which they relate. The use of this apparatus helps the linguist to observe and measure all the acoustic features other than formant structure which are operative in a complex speech signal at a given time.

The following is a summary of the conclusions reached:

(a) The acoustic displays indicated that frequency variation was general (with few exceptions) in the monosyllabic morphemes studied, even in those with very short syllables. This variation was auditorily perceived in syllables of more than 0.09 sec. duration, provided that the frequency change was sufficient to effect an auditory contrast in pitch. In syllables of less than 0.09 sec., and in those in which the frequency variation was not sufficient to effect contrast, the auditory perception was of a single pitch level.

(b) In long syllables the range of frequency was great. The auditory perception of this corresponded generally to the frequency pattern. Certain characteristic pitch patterns were observed: rising, falling, rise-fall, fall-rise, and rise-fall-rise.

(c) These patterns were communicatively significant, because they affected the meanings of the words as constituents of the grammatically structured utterances. The pattern of *Oh yes*, for example, (rise-fall-rise, fall-rise, with peak level corresponding to 300 Hz), signalled dubious assent by the speaker to the utterance of the previous speaker. The pattern of the same words *Oh yes* (rise-fall, rise-fall, with peak level 275 Hz.) signalled polite agreement with the utterance of the previous speaker.

Certain predicative phrases consisting of an intensifier and an adjective were found to have a characteristic pitch pattern. In *pretty good* a fall on *pretty* was followed

by an intra-syllabic rise on *good*. The auditory pattern agreed here with the acoustic data.

(d) Pitch appeared to be the definitive factor in auditory prominence in these morphemes. Auditory prominence corresponded well enough with frequency variation (with the modifications noted above). It often did not correspond with intensity variation. In one occurrence of *yes* (decisive agreement) a sharp contrastive frequency fall co-occurred with a sharp contrastive intensity rise.

These observations suggest that intra-syllabic pitch movement is a characteristic of spoken Australian English in certain monosyllabic words, and that it is sometimes relevant to the communication of meaning. Charting of intonation contours by using pitch levels as contour points remains a useful auditory analytical procedure. However a system of auditory analysis which concentrates upon observing levels only may not present the full facts of the language. Auditory perception does not correspond always with acoustic evidence. Instrumental displays are however most useful in helping the linguist to realize the acoustic complexity of the speech signals, and in thus elucidating the phenomena of auditory perception.

³ A recent relevant article is W. F. Klatt, "Sentence morphemes in English," *CJL*, 12, 2 (1967), 90-96.

QUELQUES REMARQUES MÉTHODOLOGIQUES

A. FRINTA

I. A mon sens la perception des sons de n'importe quel langage dépend de plusieurs circonstances. Il faut distinguer les cas de la perception;

1. chez les enfants normaux imitant leur langue du milieu (qui, à présent, n'est pas toujours maternelle),
2. chez les adultes parlant la même langue qui leur est naturelle,
3. chez ceux qui écoutent une langue étrangère qu'ils comprennent ou bien ne comprennent pas. Dans ce dernier cas la perception est moins parfaite,
4. chez les individus anormaux, soit au point de vue physiologique ou psychique.

II. La perception diffère en général, s'il s'agit des sons au commencement d'un mot ou à son intérieur: là le même son phonologique est soit soutenu par les voyelles voisines ou au contraire assimilé par les consonnes suivantes. Quant aux sons à la fin absolue des mots, leur articulation peut être affaiblie et par conséquent ils sont caduques et disparaissent. C'est ce qui est arrivé en français avec la voyelle finale *e*, dite féminin (quoiqu'elle existe aussi dans les formes verbales!) et surtout c'est le sort des consonnes finales *-s*, *-t*, *-d*, qui persistent seulement dans l'écriture étymologique, et dans la liaison avec le mot suivant.

En allemand et dans les langues slaves les consonnes finales restent fermes, si elles ne sont pas sonores, car celles-ci deviennent sourdes, par exemple en allemand Hund = Hunt, Grab/p, Weg = vĕk —, en tchèque de même: hrad, hrob, mráz, Václav — en russe les mêmes mots город, гроб, мороз, вячеслав et d'autres. Mais il y a des langues où ce changement n'a pas lieu, p. ex. en anglais: end-s = endz, dog-s, où les consonnes finales sont perçues sous leur forme sonore. De même dans les cas, si elles sont suivies de la voyelle *e* muet, p. ex. mov(e), = mŭv, rose = rouz, ce qui est commun aussi en français: cav(e), mod(e), chos(e) = šōz) etc.

III. C'est aussi l'*accentuation* qui peut y jouer son rôle. J'ai déjà mentionné la disparition de la voyelle *-e* atone, qui arrive en français non seulement à la fin, mais aussi à l'intérieur des mots, p. ex. vie, rare; prom(e)nade. Un cas spécial de la voyelle initiale atone présente le mot grec apothéké qui est devenu boutique en français, bien sûr par la combinaison avec l'article: l'aboutique > la b.

IV. Il est très instructif de suivre les *changements* phonétiques qui ont lieu dans les mots *empruntés* d'une autre langue, ordinairement voisine. C'est donc la phoné-

tique historique qui nous fournit des éclaircissements sur le caractère des sons perçus de la langue originale et imités par l'ouïe.

V. C'est surtout de *l'articulation* que dépend la perception: les consonnes *b d g* ne sont pas, en allemand, sonores dès leur début (comme elles le sont dans les langues romanes, slaves et en anglais).

C'est pourquoi elles sont perçues au début des mots plutôt comme leurs correspondantes *p t k*, p. ex. Bürgermeister (autrefois Burgmeister) a donné en tchèque purkmistr, Bärenstein > Per(n)štejn (château en Moravie). Seulement à l'intérieur des mots, appuyé par une autre consonne sonore, *b* persiste: Nymburk (ville en Bohême). Un exemple pour *d* initial: allemand doppel (pour français double) > en tchèque populaire tupl-ovaný; pour *g* initial: allemand Grobian > en tchèque vulgaire krobián, gar > kór.

VI. Si la voyelle allemande *ü* n'est pas assez arrondie par les lèvres, elle devient *i*, même dans les patois allemands et surtout dans l'usage des Slaves: Kurfürst kurfürť en tchèque, Müller > Miller = nom de famille tchèque, polonais, russe.

DISCUSSION

Martens:

Il est toujours utile si l'on peut faire voir aux étudiants la différence acoustique. De ce point de vue-là il se montre agréable d'avoir des mots tchèques dans lesquels *b*-assourdi de „Bürgermeister“ par perception tchèque devient un *p* sourd „purkmistr“. Mais il faut se méfier, car il y a aussi d'autres cas comme: allem. „Torte“ (avec *t* sourd) tch. „dort“ ou comme allem. „Decke“ — tch. „deka“ etc.

REACTION-TIME EXPERIMENTS IN THE STUDY OF SPEECH PROCESSING

D. B. FRY*

There is now a good deal of information available about the acoustic cues which are used in speech and their relation to the phonological system in certain languages. This knowledge is on the whole confined to the operation of single cues and to the initial stages of speech processing in the reception of speech. We understand in fact very little about the way in which acoustic cues for recognition are combined together and even less about the way in which the first stages of recognition are linked with the succeeding operations of linguistic processing.

The experiments described in this paper represent simply a first attempt to advance the enquiry by dealing with more complex situations and they are concerned more with finding a means of doing so than with producing far-reaching results.

First let us summarize very briefly the sequence of operations we may expect to take place in the reception of a spoken message: the acoustic input of speech is converted by the receptors into perceptual patterns which present a complex of features; these features are correlated with the acoustic cues and the listener has learned to make use of various combinations of cues in recognizing the sounds; the incoming sounds are assigned to the phonemic categories of the language on the basis of long-term *a priori* knowledge of the categories and short-term sequential information; the phonemic string forms the input for successive stages of further linguistic processing which yield morphemes and words which make up the message.

For the sake of convenience we can divide these operations into two main parts and refer to all the processing which depends directly on the acoustic cues and their combination as primary recognition, and the subsequent stages as linguistic processing. Either of these may call for operations of greater or less complexity. There will be contexts in which primary recognition is a simple operation, perhaps depending on the evaluation of a single acoustic cue; there will be others in which it is very much more complex, where it will be necessary for the listener not only to deal with a number of cues for a single phoneme recognition, but to process cues for a phoneme sequence where acoustic cues are interdependent and perhaps to operate at the same time upon cues for certain prosodic features. Similarly in linguistic processing, the

* University College London.

recognition of a phoneme string as a single-morpheme word may call for a comparatively simple piece of processing, whereas the string which forms a polymorphemic word, involving syntactic rules of a complicated kind, will require much more complex processing.

Although we at present know very little about the nature of these processing operations, it would be an error to assume that they are necessarily done serially in time. In fact, the most conspicuous feature of the functioning of the human brain is its very great capacity for doing many things at the same time; it seems able to employ the most intricate patterns of parallel working in such a way as to try out and discard many solutions to a problem in a very short time and hence to discover short cuts to correct solutions. Nonetheless, in the particular case of speech reception, it seems intuitively necessary that a complex piece of processing should take longer than a single simple operation. This is partly because the speech input is necessarily strung out in time, and a complex operation is likely to depend on information spread over a longer stretch of the acoustic continuum, and partly because of the hierarchical nature of language systems which requires that decisions on a lower linguistic level be made before processing on a higher level can be completed.

If it is the case that the more complex the speech processing, the longer it takes a listener to complete it, and if we could find some reliable means of determining the time that is required, we should have a way of gaining at least a qualitative idea of the complexity of the processing needed in a given case and perhaps eventually a criterion for distinguishing different processing operations. It is with this purpose in view that these reaction-time experiments have been begun. By setting a listener a variety of speech processing tasks and providing him with a means of signalling when he has completed the task, we may begin to gather evidence of the kind we have just referred to.

It is necessary first of all to establish whether the processing time does in fact appear to increase with the complexity of the task, and this can be done by beginning with very simple operations at the level of primary recognition, such as asking a listener to distinguish between words forming a minimal pair in his native language. The experiments reported in this paper do not in fact go much further than the exploration of this stage of the problem and the indication of some directions in which further progress seems likely.

THE TECHNIQUE OF REACTION-TIME EXPERIMENTS

At this point, it is necessary to say something briefly about the nature and the technique of reaction-time experiments. The essence of the method, as we have said, is to set the experimental subject some task, in this case a speech reception task, and to get him to signal, by pressing a button or a key, or perhaps by speaking, the moment when he has completed the task. The time interval between the arrival

of the speech stimulus at his ear and the making of the response is the reaction-time. In these experiments, subjects responded by pressing a key. The method of asking the subject to speak back as a response was rejected because the object of the work was to investigate the operation of the speech reception mechanism, which we cannot suppose to be independent of the speech generating mechanism. To ask subjects to speak a response would be to set the speech generating mechanism working at the same time as the reception mechanism which is under observation. It seemed preferable to make the response a motor act unconnected with the speech mechanism and thus avoid the problem of determining from the data what part was played by the speech generating, as distinct from the speech reception, mechanism.

To take the simplest of the experiments as an example, the subject is asked to listen to words which are fed to him through telephone receivers; the words are in recorded natural speech, reproduced in good listening conditions so that they are easily recognized. A particular test, for example, consists of the words [bit] and [bet], occurring in random order; the subject has two keys, one marked *bit* and the other *bet* and his task is to press the appropriate key as soon as he has decided which word he has heard. This arrangement can be viewed schematically as it is shown in Fig. 1. The acoustic input to the two ears, suitably transformed by the hearing mechanism, provides the basis for the primary recognition of the sequence and this is followed by whatever linguistic processing may be necessary. This leads in turn to the decision as to which key is to be pressed and thus initiates the motor response, that is the neural command and the muscle action in the arm and finger which presses the key. When the speech input is changed, the operations represented in the left-hand side of the diagram remain essentially unchanged.

It must be noted, however, that reaction-times, in almost all circumstances, show considerable variability, even within one individual subject. If a subject is asked not even to make a choice but simply to press a key each time he hears a click in his

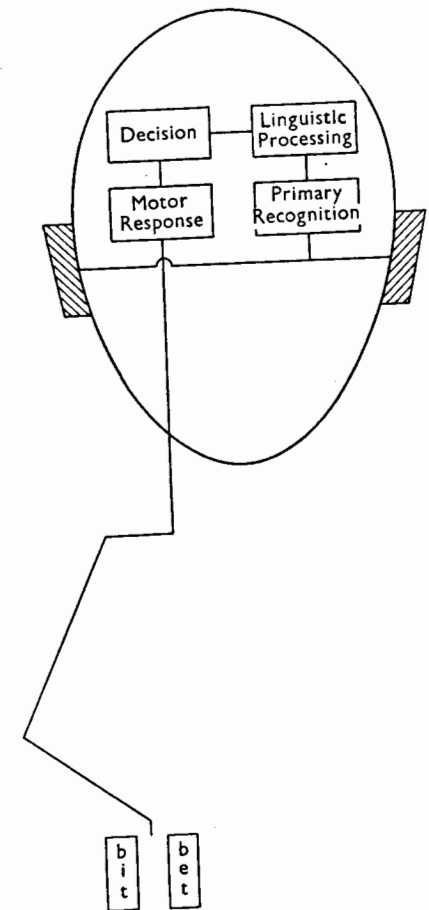


Fig. 1. Schematic diagram of response circuits involved in reaction-time experiments.

telephones, his response time (the simple reaction-time) will vary from moment to moment, from hour to hour and from day to day. That is to say that the operations represented by the left-hand side of the schematic diagram of Fig. 1 will make a contribution to the total reaction-time which is variable. In the present experiments the effect of this variability was minimised in the following way. In a given test-run, the subject was asked only to distinguish between two familiar words which were clearly audible to him, e.g. [bit] and [bet]. He was told to listen to each test item and as soon as he was certain which word had been spoken, he was to press the key marked with that word. He would then respond to a random sequence of the words [bit] and [bet] until he had made 50 responses to each word. This test-run takes about five minutes and whatever variations take place in the response mechanism in this short interval will tend to affect the responses to [bit] and [bet] equally, since they occur randomly during this period. Statistically significant differences in the reaction-times will consequently be due mainly to differences in the speech processing. This method of direct comparison was adopted throughout these experiments.

INDIVIDUAL DIFFERENCES IN SPEECH PROCESSING

A further source of variability, and for the purposes of these experiments a more important one, is to be found in the linguistic behaviour of individual listeners. Although the use of language for communication purposes depends upon the use of a common system, there is likely to be considerable individual variation certainly in the speed of processing and also to some degree in the sequence of operations. In decoding a spoken message, the listener's task is to arrive at the correct, that is the commonly accepted, solution; he retains a certain freedom as to the steps by which he reaches the solution and possible variations of this kind are important to our understanding of speech processing. It would therefore be a mistake to deal with the problem of individual variation by pooling reaction-times from a large number of subjects since by doing so we should lose some of the qualitative information that is of most interest to us. All the results given in this paper are therefore measurements from individual listeners who have made each a large number of responses to the test material.

REFERENCE POINTS FOR REACTION-TIME MEASUREMENTS

The first experiments to be described deal with very simple and basic considerations which are intuitively well understood. The technique is to ensure that the listener is in no doubt about the words he is going to hear, and to ask him to press the appropriate key as soon as he possibly can, but of course without making mistakes. If he is dealing with minimal pairs, the moment at which he can make a decision will clearly depend upon the point in the word at which the minimal

difference occurs. Transposing this into terms of primary recognition, at some point in the input the listener will pick out some acoustic cue or cues which will identify the member of the pair of words. Reaction-time might therefore be defined as the time interval between this moment and the subject's response. In practice we cannot be sure what cues the listener will use for the purpose nor at what moment they will appear in the input. For convenience, therefore, it is preferable to adopt some point of reference which is more easily defined and in all the measurements given here this point is the beginning of the word as determined from an oscillogram. This means the noise burst of a plosive sound, the beginning of friction noise in a fricative, the first voice cycle in the case of a voiced continuant, and so on.

If the reaction-time is then taken as the time interval between the beginning of the word and the pressing of the response key, we shall expect that this time will increase progressively as the minimal distinction between pairs of words occurs later and later in the words. This is in fact the case. If we take the word [bit] and place it in contrast first with the word [pit], then with [bet] and then with [bid], for each individual listener the reaction-time becomes progressively longer. This effect is presented in Fig. 2 where the spectrogram is that of the utterance [bit] used as the stimulus in the tests. The arrows indicate mean response times and all of them refer to the time taken by the subject to press the *bit* key, not to the time taken to respond to the contrasted words. Reaction-time is measured from the burst of the [b], and in these results, for subject N, the mean time taken to press the response key for [bit] when it is contrasted with [pit] is 325 msec., when contrasted with [bet], 359 msec. and when contrasted with [bid], 431 msec. Each value is the mean of 50 responses by the subject; the difference between the first two means is significant at a probability level less than 0.05, the difference between the second and third means is highly significant, at a probability level well below 0.001.

This effect, as we should expect, continues to appear when the sequence is extended to disyllables or trisyllables. Fig. 3 shows the mean reaction-times for the same subject responding to the contrast [big] — [bid], [bigin] — [bigan] and [biginɪŋ] — [biginə]. The means for [big], [bigin] and [biginɪŋ] are 430 msec., 522 msec. and 630 msec. respectively, differences between successive means being highly significant with a probability value well below 0.001.

These simple tests show one interesting fact, and that is that in the conditions of the experiment, subjects have no difficulty whatever in responding before a word or syllable is complete; the processing is capable of dealing with segments smaller than the whole word or syllable. This emerges more clearly if we normalise the results by referring the reaction-time in each case to the duration of the stimulus. Table 1 shows the means for subject D for the same set of contrasts, with reaction-time expressed as a proportion of the total duration of the stimulus. In only three cases [big], [bid] and [biginə], does the mean reaction-time exceed the total duration of the stimulus, and even here two of the values are almost equal to the duration.

Table 1. Subject D.

Mean Reaction-Time divided by Stimulus Duration				
[bit]	0.666	contrasted with	[pit]	0.750
[bit]	0.748	[bet]	0.734
[bit]	0.983	[bid]	1.084
[bid]	0.930	[big]	1.224
[bigin]	0.839	[bigan]	0.748
[biginin]	0.906	[biginə]	1.030

Table 1. Mean reaction-times for an individual subject expressed as a proportion of the duration of the stimulus word.

PROCESSING OF CONSONANT CLUSTERS

We will now turn to tasks of a more complex kind, which still depend on the processing of acoustic cues and very little more. In these experiments subjects responded in the way already described to a whole series of contrasted words in successive tests. The list of these contrasts is given in Table 2, together with the mean reaction-times for subject C. It will be seen that the series builds up from minimal pairs with single initial consonants, to those with initial three-consonant clusters.

In the results that have been given so far we have been mainly concerned with differences in the reaction-time for the same word when it is placed in different contrasts. It is also possible that, when two words are contrasted, one of the two may produce reaction-times which are significantly shorter than those for the other. If this happens, it is in a general way a good indication that by this technique we are getting at differences in speech processing since in a single test other factors which might account for differences are balanced out.

Table 2 shows that in this series of tests there were six contrasts in which this subject gave significantly shorter reaction-times for one word than for the other. For example, when [plei] was contrasted with [lei], the mean was 346 msec. for [plei] and 394 msec. for [lei]. It was in fact generally true that the shortest reaction-times were provided by words beginning with a plosive when contrasted with words not beginning with a plosive. The sharp wave-front of the plosive onset is evidently a cue which is very rapidly processed. It is interesting, however, that in a contrast such as [plei] — [lei] the situation is asymmetrical in the sense that presence of the plosive wave-front is quickly registered yet its absence cannot be used as a cue; the reaction to [lei] has to wait for positive indications that the sound is [l]. The reaction-time for [lei] contrasted with [plei] is about equal to that for [lei] contrasted with [rei]. Since the principal acoustic cue for [l] and [r] is found in a rather slow frequency change in the second and third formants, it is not surprising that the reaction-times in these cases are rather long.

Table 2. Subject C.

Mean of Reaction-times (msec.)				Mean of Reaction-times (msec.)	
[rei]	417	contrasted with	[lei]	405	
[rei]	387	[prei]	370	
[rei]	346	[sprei]	372*	
[prei]	332	[sprei]	353***	
[prei]	490	[plei]	494	
[sprei]	507	[strei]	487	
[sprei]	566	[splei]	637***	
[lei]	393	[plei]	346***	
[lei]	379	[slei]	381	
[lei]	365	[splei]	385	
[plei]	323	[slei]	347***	
[plei]	302	[splei]	367***	

* Significant difference between means within contrast, $p < 0.05$.
 *** $p < 0.001$.

Table 2. A series of contrasts involving single initial consonants and initial consonant clusters with mean reaction-times for an individual subject.

Another interesting case is that of contrasts in which one word begins with friction noise and the other does not. In the case of [plei] versus [splei], we have for [plei] the shortest reaction-time in the whole series, 302 msec., which may well be the combined effect of the absence of friction and the presence of plosive onset. There is some tendency for the situation to be asymmetrical but with the absence of friction giving shorter times than its presence, although the differences do not always reach significance.

The most striking effect in this series of experiments, however, was seen in cases where consonant clusters of similar structure were contrasted with each other. In the examples we have just been discussing, we can see that the listener could arrive at a decision by giving his attention to a single acoustic cue—the sharp wave-front, the friction noise or the third formant frequency change. When two- and three-consonant clusters are contrasted with each other the situation seems to be much more complex. The contrasts [prei] — [plei], and [sprei] — [strei], [sprei] — [splei] produce a fairly dramatic increase in the time required for the task. Fig. 4 presents some of the relevant results in a somewhat clearer way. The spectrograms are analyses of two of the stimuli used in the tests, the left-hand one [sprei] and the right-hand,

[*splei*]. The arrows show on the time-scale the mean reaction-time for [*sprei*] and [*splei*] when they are contrasted with the word shown at each arrow. The mean reaction-time is 353 msec. when [*sprei*] is contrasted with [*prei*], and 372 msec. when it is contrasted with [*rei*]. We are here looking at means from different test-runs and there is no statistical significance in this difference. But when [*sprei*] is contrasted with [*strei*], the mean is 507 msec. and when it is contrasted with [*splei*], 556 msec. The difference between these values and the first two means is very highly significant, well beyond the level required for a probability of 0.001. The situation is paralleled when we look at the time taken to press the key for [*splei*] in different contrasts: in the [*plei*] and [*lei*] contrasts, the means for [*splei*] are 367 msec. and 385 msec., yet with [*sprei*] the mean is 637 msec.

The effect is equally marked in the case of the two clusters [*pr-*] and [*pl-*], shown in Fig. 5. The contrast [*prei*] — [*plei*] gives very long reaction-times, despite the fact that the other contrasts, with [*sprei*], [*rei*], [*splei*] and [*lei*], give some of the shortest reaction-times. Again the differences between the long and the short reaction-times are statistically highly significant.

It is very difficult to explain this effect except on the grounds that more complex processing is required in these cases. There is no doubt that the same acoustic cues are available in the various contrasts, and we can only conclude that the listener uses the minimum of cues that will enable him to make a decision in any given conditions. The processing is reduced to the simplest possible operations and when this means contrasting, for example, presence and absence of friction noise, the time required is very short. When the contrast is between [*pr-*] and [*pl-*], between [*spr-*], [*str-*] and [*spl-*], however, it seems that more acoustic cues have to be taken into account, the processing is more complex and the time required is correspondingly long. Although these clusters were all in initial position in the syllable, the mean reaction-times were almost invariably in excess of those needed for the contrast of final consonants in [*bit*] and [*bid*], and this despite the fact that the duration of the stimulus word [*prei*] was identical with that of the word [*bit*].

PROCESSING OF MORPHEME BOUNDARIES

The experiments already described have all been confined to the processing of acoustic cues, with no demand for subsequent linguistic processing. It would clearly be very valuable if the reaction-time technique could be used to explore higher levels of processing. This last section deals with one set of preliminary experiments carried out to see whether there are possibilities in this direction.

These tests involved three contrasts in which a morpheme boundary occurred. [*pik*] — [*piks*], [*pik*] — [*pikt*], [*piks*] — [*pikt*]. Several factors make it difficult to draw firm conclusions from these tests, apart from the fact that many more contrasts need to be investigated. In the nature of the English language system, such morpheme boundaries involve word-final consonants or consonant clusters so that

the acoustic cues on which the processing is based come towards the end of the word and reaction-times measured from the beginning of the word will in any case appear long. Furthermore, it is difficult if not impossible to find in English minimally contrasted sequences of which one involves a morpheme boundary and the other does not without having recourse to orthographic differences which are doubtfully represented at the phonetic level.

A typical set of mean reaction-times for the [*pik*] contrasts is given in Fig. 6. It is true that all the mean reaction-times are greater than the duration of the stimulus word, sometimes by a considerable amount, but this is not unexpected in view of the position of the contrasting sounds. One might guess that if listeners were operating on the very simple acoustic basis which was noted in the case of [*prei*] — [*sprei*], etc., the response might well come sooner than it does here. On the evidence of this one set of contrasts, however, it is impossible to say whether the presence of the morpheme boundary is affecting the processing time or not.

This question along with many others, including the whole field of the processing of prosodic features, may be answered by further experiments along these lines. The preliminary results reported in this paper do at least suggest that there is much to be learned about speech processing by measuring how long it takes listeners to carry out a particular set of operations.

DISCUSSION

Hill:

I do not wish to question, but to say why I feel this one of the most significant phonetic papers I have heard in years. We have been hearing lately a great number of statements that processing of what we hear may not be linear, but all at once even reversed. This is the first time I have heard anyone deal with the problem directly.

Lehiste:

Was there any learning effect observed in the responses?

Newell:

Would not the fact that you were using only one acoustic waveform for each stimulus, rather than different attempts at the same words by the same speaker, decrease the reaction time, due to recognition of slight differences in the phonetically similar parts of the stimulus.

Fry:

Ad Lehiste: I have not examined systematically the statistical variation in reaction-time within each test, but from mere inspection it seems that if there is any degree of learning in making the responses it takes place very rapidly indeed. The subject received no warning signal when the first stimulus in a test was about to arrive, and for this reason the very first reaction-time was sometimes quite long, but apart from this, it was never possible to note by inspection that the first few times were longer than succeeding ones. Furthermore, there were two test runs for

each minimal pair and there was no tendency for the mean of the first 25 responses to a given word to be greater than the mean of the second 25. This did sometimes happen but about equally frequently the second was greater than the first. Had there been a pronounced learning effect lasting over some minutes, there would have been a systematic trend towards shorter reaction-times in the second of each pair of test runs.

Ad Newell: While I think it is possible that if the amount of variation in the stimulus words were increased, the reaction-time might increase. I do not think that this is beyond question. If by this reaction-time method we are really able to get a measure of the acoustic-linguistic processing time, then it is conceivable that the introduction of a certain degree of variation in the stimulus may not lead to a significant increase in reaction-time.

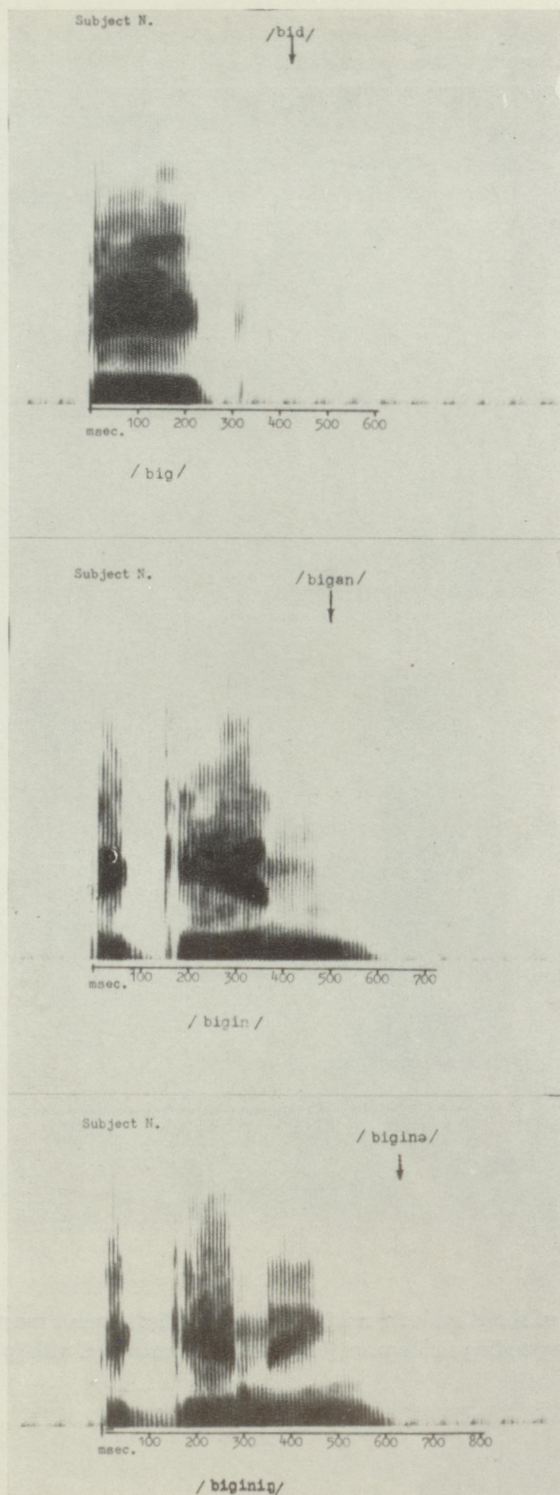


Fig. 3. Spectrograms of the stimulus words *[big]*, *[begin]* and *[beginiŋ]*. Arrows indicate mean reaction-times for an individual subject responding to the words *[big]*, *[begin]* and *[beginiŋ]* when contrasted with *[bid]*, *[bigan]* and *[biginə]* respectively.

Fry: Reaction-Time Experiments in the Study of Speech Processing

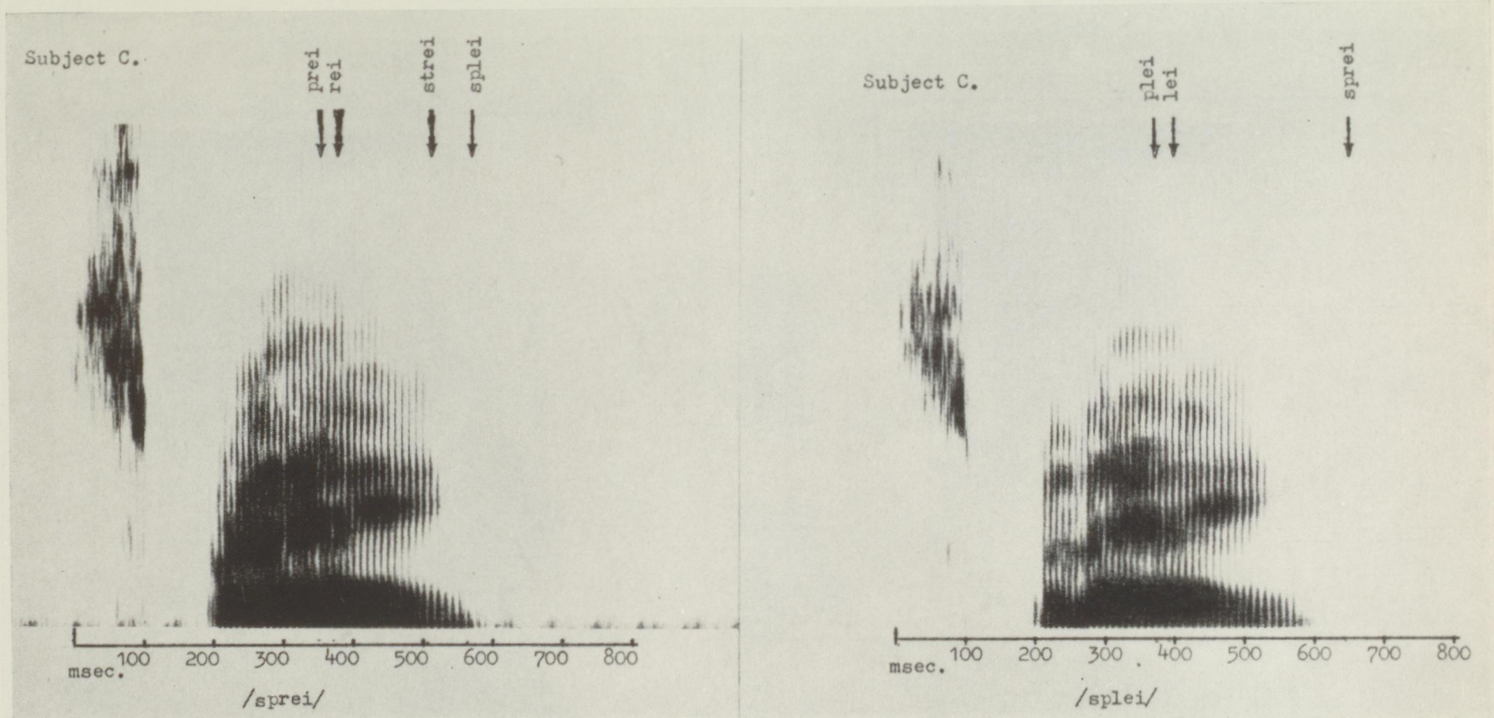


Fig. 4. Spectrograms of the stimulus words [sprei] and [splei]. Arrows indicate mean reaction-times for an individual subject responding to these two word in the contrasts [sprei] — [prei], [sprei] — [rei], [sprei] — [strei], [sprei] — [splei], and [splei] — [plei], [sprei] — [lei] and [splei] — [sprei].

Fry: Reaction-Time Experiments in the Study of Speech Processing

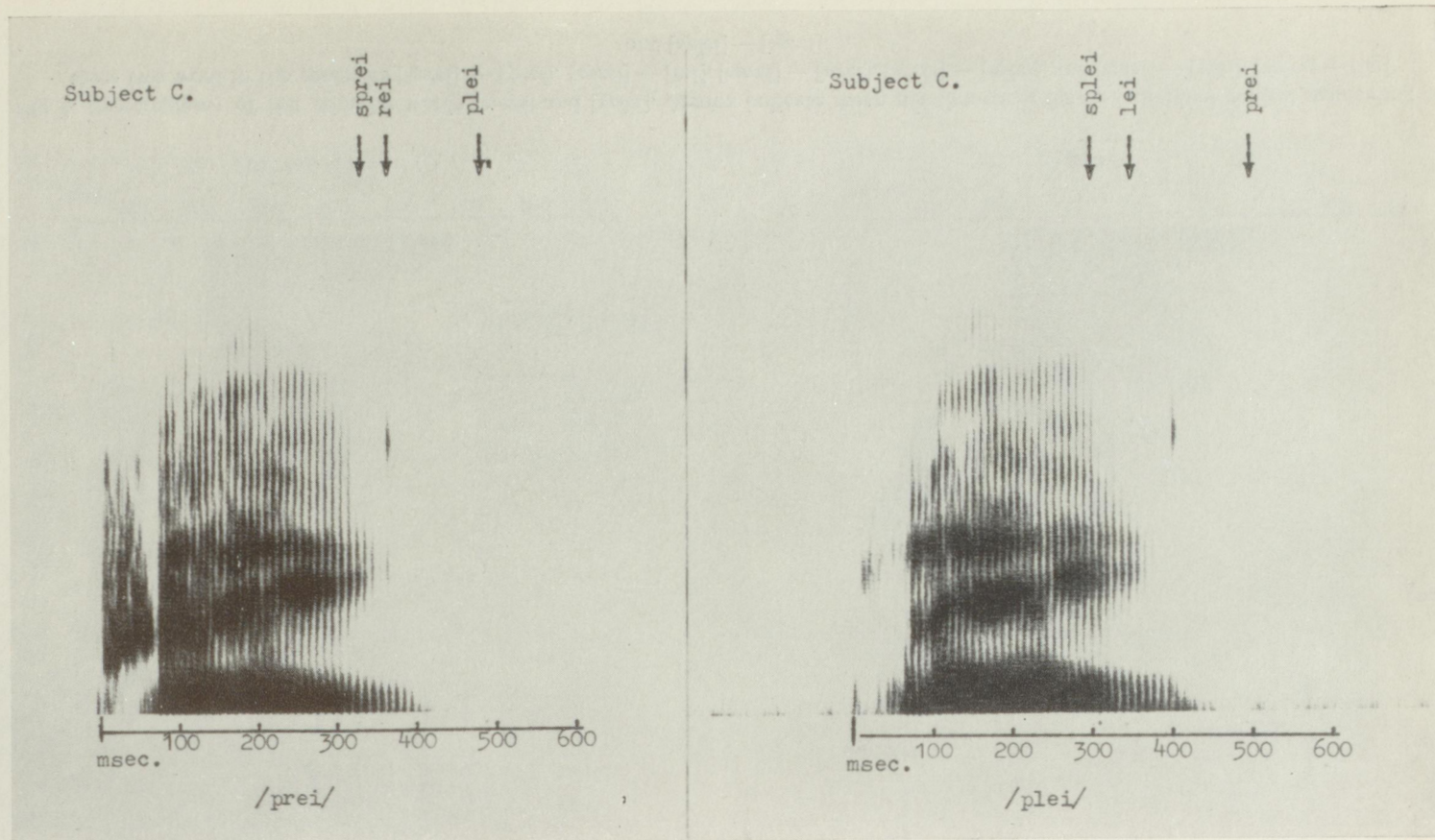


Fig. 5. Spectrograms of the stimulus words [prei] and [plei]. Arrows indicate mean reaction-times for an individual subject responding to these two words in the contrasts [prei] — [sprei], [prei] — [rei], [prei] — [plei], and [plei] — [splei], [plei] — [lei] and [plei] — [prei].

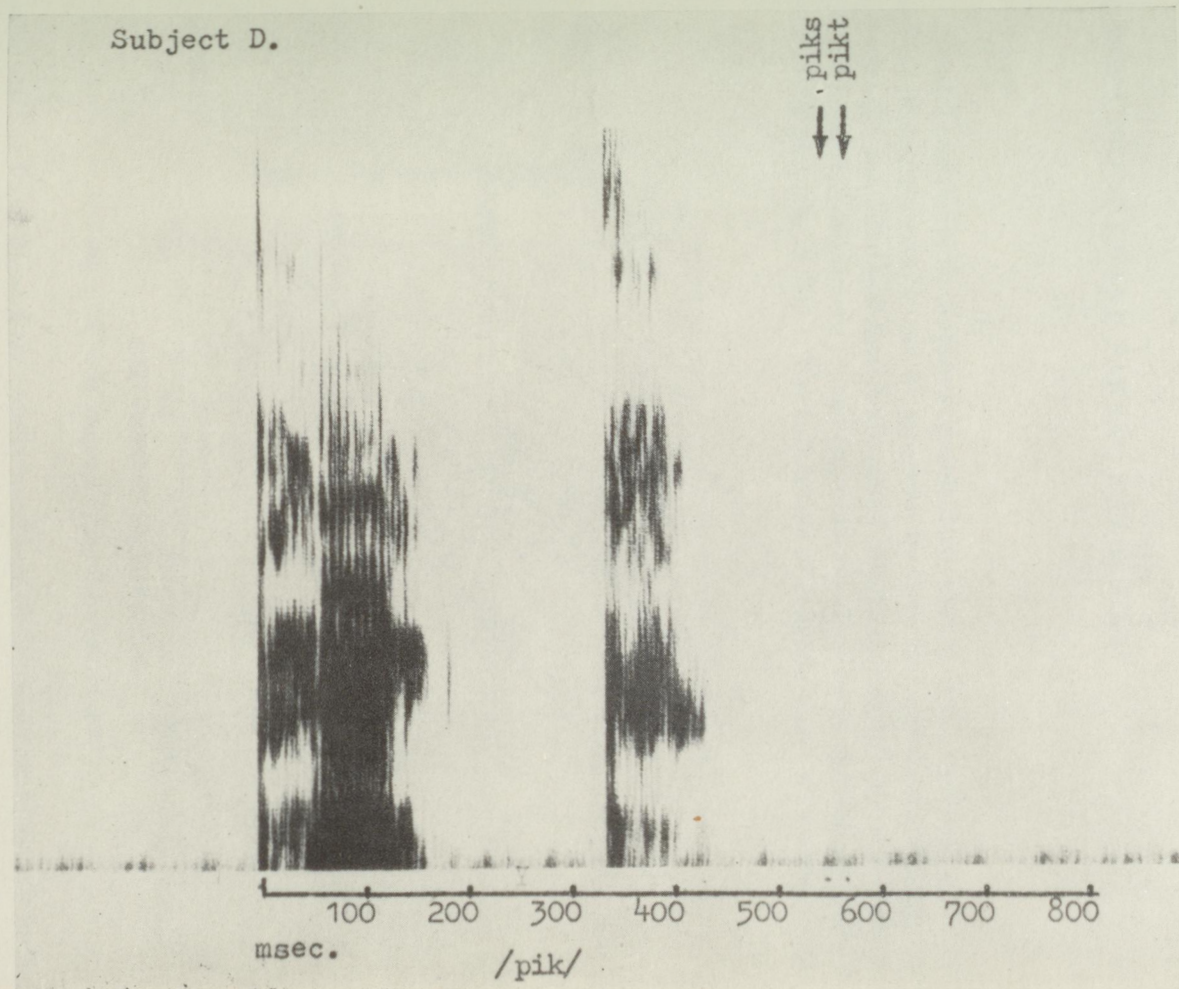


Fig. 6. Spectrogram of the stimulus word [pik]. Arrows indicate mean reaction-times for an individual subject responding to this word contrasted with the verb forms [piks] and [pikt].

ROLES OF PITCH AND HIGHER FORMANTS IN PERCEPTION OF VOWELS

HIROYA FUJISAKI AND TAKAKO KAWASHIMA*

Spectral analysis of the Japanese vowels shows that the five vowels [a], [e], [i], [o] and [u] of a single speaker can be separated by their first and second formant frequencies (F1 and F2). Considerable amount of overlap is observed, however, between [u] and [e] as well as between [o] and [a], when vowels of many speakers are plotted on F1—F2 plane, as in Fig. 1. The two vowels in each pair share approximately the

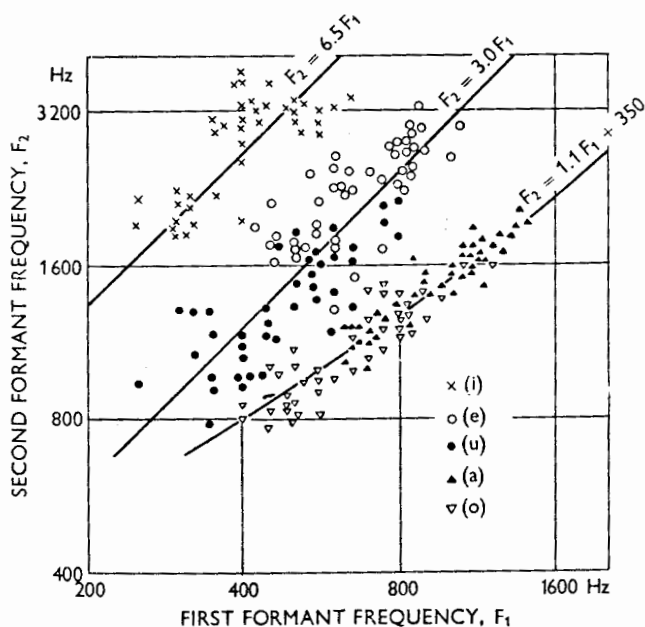


Fig. 1. F₁ — F₂ diagram of the Japanese vowels.

same ratio of F2 to F1, and the overlap can be ascribed mainly to differences in the length of vocal tract. A normalizing process, based presumably on higher formant frequencies (F3, F4 etc.), is expected in the identification of these vowels. It is not

* Department of Electronic Engineering, University of Tokyo.

clear, however, whether concurrent changes of pitch frequency (F0) and formant frequencies, as observed in Fig. 2, are necessary in the normalizing process. In order to estimate the importance of these parameters as perceptual cues, the following method has been adopted using synthetic vowels.

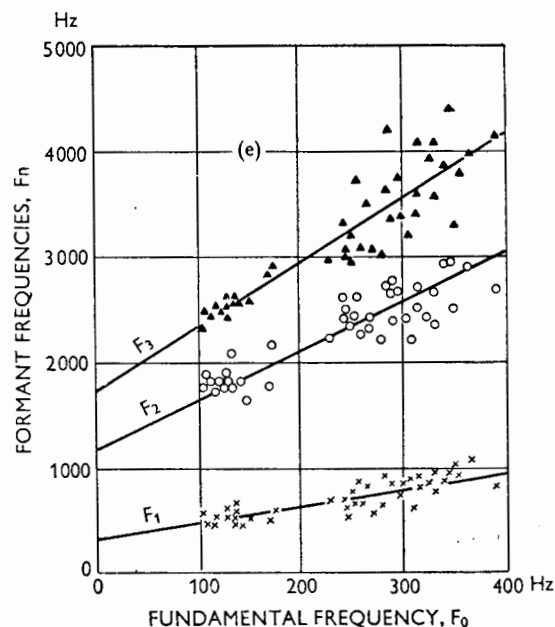


Fig. 2. Fundamental frequency F_0 versus formant frequencies F_1 , F_2 and F_3 for [e].

Method. The frequency parameters of synthetic vowels generated by the terminal analog synthesizer of Fig. 3 are F_0 , F_1 , F_2 , F_3 , F_4 and Kr4 (higher pole correction), all of which are variable. The spectral envelope of buzz source has no zeros and falls with a slope of -12 dB/oct. from 170 Hz. When F_1 and F_2 are varied along the direction of overlap, (for example, $F_2/F_1 = 3$ for [u] and [e]), keeping all other parameters constant, a perceptual transition is observed between the two vowels (Fig. 4), and the values of F_1 and F_2 corresponding to 50% responses can be defined as the boundary values F_{1b} and F_{2b} .

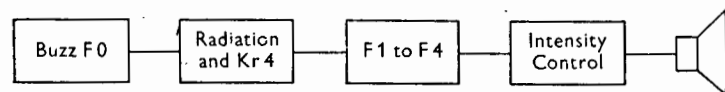


Fig. 3. Block diagram of the vowel synthesizer.

In the following experiments, the effects of i) changes of F_3 , F_4 and Kr4 as a group, ii) change of F_0 only, and iii) concurrent changes of F_0 , F_3 , F_4 and Kr4 were examined. In case iii), an experimental relationship $F_3 = 6.2 (F_0 + 270)$ Hz was adopted, and F_3 , F_4 and Kr4 were varied always following the pattern for the neutral vowel.

The ranges for F_0 , F_3 and F_4 were 140 — 345 Hz, 2.5—4.0 kHz, and 3.5—5.6 kHz, respectively. The duration of the stimuli was 0.8 sec and a rise time constant of 20 msec was given to pitch and intensity for naturalness. For the purpose of comparison, the effects of higher formant changes were also examined for vowel sounds excited by a noise source which had a flat spectrum up to 8 kHz.

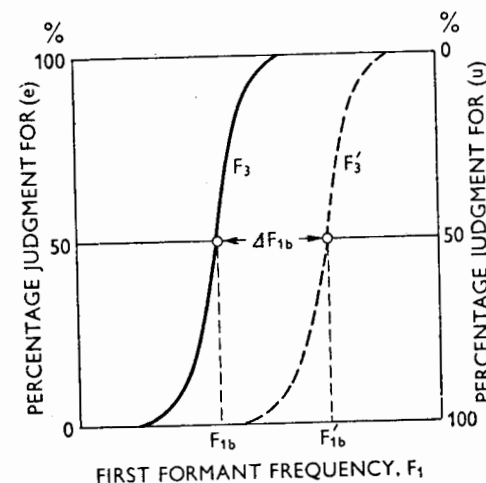


Fig. 4. Perceptual transition between [u] and [e] and the boundary F_{1b} as affected by a change in F_3 .

Results. Table 1 summarizes the results of listening tests performed on 14 subjects (10 adults and 4 children). Since no significant differences were observed due to age or sex, mean values for all subjects are listed. The results show that, for [u] — [e] identification, the rôles of higher formants are not so important as those of

Table 1. Influence of F_0 , F_3 , F_4 and Kr4 on the perceptual boundaries between [u]—[e] and [o]—[a].

Source		[u]—[e]	[o]—[a]
Buzz	i) $\frac{\Delta F_{1b}}{F_{1b}} \quad \frac{\Delta F_3}{F_3}$	0.63	0.13
	ii) $\frac{\Delta F_{1b}}{F_{1b}} \quad \frac{\Delta F_0}{F_0}$	0.22	0.29
	iii) $\frac{\Delta F_{1b}}{F_{1b}} \quad \frac{\Delta F_3}{F_3}$	1.00	0.93
Noise	$\frac{\Delta F_{1b}}{F_{1b}} \quad \frac{\Delta F_3}{F_3}$	0.96	0.76

Table 1. Influences of F_0 , F_3 , F_4 and Kr4 on the perceptual boundaries between [u] — [e] and [o] — [a].

F1 and F2, and concurrent changes of pitch and higher formants are necessary to counteract the changes in F1 and F2. For [o] — [a] identification, the tendency is still greater and the influences of higher formants are even less than that of pitch alone. In noise-excited vowels, on the other hand, the rôles of higher formants are as important as the combined rôles of pitch and higher formants in buzz-excited vowels.

Conclusion. A method has been described to examine the rôles of pitch and higher formants in the perception of vowels. Results of experiments have revealed that, for ordinary buzz-excited vowels, perceptual normalization is not complete unless pitch and higher formants vary concurrently.

ВОПРОСЫ ТЕОРИИ ДИФФЕРЕНЦИАЛЬНЫХ ПРИЗНАКОВ

Ш. Г. ГАПРИНДАШВИЛИ*

Рассмотрение фонемы как пучка дифференциальных признаков подразумевает предварительное определение фонетических характеристик последних.

Дихотомическая классификация дифференциальных признаков фонем страдает отсутствием единого критерия классификации и искусственностью членения на цепочки бинарных противопоставлений более сложных и многочленных противопоставлений, отражающих сложную природу фонологических систем языков.

„Бинаризм оказывается скорее умозрительным построением, чем попыткой согласовать между собою результаты . . . наблюдений“ — (А. Мартин). Терминология теории дифференциальных признаков . . . „не отвечает системным отношениям и иерархии единиц признаков любой данной фонематической системы“ (А. Реформацкий).

Речь — процесс фильтрации, видоизменения динамики и спектральной структуры несущего потока. Поэтому, при классификации дифференциальных признаков фонем, мы должны учитывать признаки характера, направления и динамики несущего потока наравне с тембровыми признаками и признаками акустических градаций.

1. ПРИЗНАКИ ХАРАКТЕРА, НАПРАВЛЕНИЯ И ДИНАМИКИ НЕСУЩЕГО ПОТОКА

1. А. Голосовые

В. Смешанные

С. Шумные

1. Большая общезвуковая интенсивность; квазипериодические колебания на всем протяжении артикуляции; концентрация звуковой энергии в диапазоне низких частот спектра.

1. В. Одновременное наличие квазипериодических низкочастотных и аperiodических высокочастотных колебаний; распределение звуковой энергии по большому диапазону спектра.

* Институт языковедения АН Грузинской ССР, Тбилиси.

1. С. Высокочастотные аperiodические колебания на всем протяжении артикуляции или чередование их с безголосым и бесшумным интервалом; концентрация звуковой энергии в диапазоне высоких частот спектра.

Класс 1. А. включает гласные и сонорные фонемы. В класс 1. В. входят звонкие щелинные и аффрикаты; в класс же 1. С. — глухие щелинные и смычно-взрывные, в том числе придыхательные и смычногортанные мгновенные и аффрикаты.

2. А. Непрерывные В. Пульсирующие С. Прерванные

1. А. Равномерное распределение (при почти стационарной формантной структуре) общезвуковой интенсивности на всем протяжении артикуляции.

2. В. Многократное нарастание и спад интенсивности с незначительными изменениями и в формантной структуре.

2. С. Неравномерное распределение общезвуковой энергии по отдельным фазам артикуляции; резкие перепады интенсивности и значительные изменения формантной структуры в разных фазах артикуляции.

В класс 2. А. попадают гласные, сонорные, за исключением выбрантов, и щелинные согласные. В класс 2. В. — выбранты: *P, γ, X*. в класс 2. С. — же — все смычно-взрывные, как мгновенные, так и аффрикаты.

3. А. Звонкие В. Придыхательные С. Смычногортанные

3. А. Квазипериодические колебания во всех фазах (позиционно : в информационном отношении наиболее нагруженной фазе) артикуляции; обязательное наличие области усиления в диапазоне низких частот спектра в указанных выше фазах артикуляции.

3. В. Заполнение высокочастотными аperiodическими колебаниями интервала между взрывом данной и началом последующей фонемы; концентрация звуковой энергии, преимущественно, в диапазоне высоких частот спектра.

3. С. Наличие ничем незаполненного, безголосового и бесшумного интервала между взрывом данной и началом последующей фонемы.

В разряд 3. А. находят место гласные, сонорные, а также смешанные, характеризующиеся наличием периодических и аperiodических колебаний одновременно, или в масштабе фонемы. Класс 3. В. включает в себя глухие придыхательные смычно-взрывные и аффрикаты в языках с фонологическим противопоставлением : придыхательный — непридыхательный. Класс 3. С. объединяет смычногортанные мгновенные и аффрикаты.

4. А. Оральные В. Назализованные С. Назальные

4. А. Отсутствие прохода для фонационного потока в носовую полость; отсутствие областей усиления в диапазонах частот, соответствующих носовому резонансу назализованных; наличие областей усиления в диапазонах частот, соответствующих диапазонам антирезонанса назальных.

4. В. Одновременное включение в артикуляцию резонаторов рта и носа; наличие в спектре области усиления, соответствующей диапазону носового резонанса.

4. С. Участие в артикуляции носового резонатора при параллельном подключении закрытого ротового резонатора; наличие в спектре нулей, соответствующих диапазону ротового резонанса.

5. А. Сильные В. Простые С. Слабые

5. А. Большая общезвуковая интенсивность в информационном отношении в наиболее нагруженной фазе артикуляции; напряженность артикуляционных органов; более отчетливая формантная структура по сравнению с простыми и со слабыми.

5. В. Средняя общезвуковая интенсивность и такая же степень напряженности артикуляционных органов; средняя отчетливость формантной структуры.

5. С. Малая общезвуковая интенсивность; отсутствие напряженности артикулирующих органов; бедность и нерельефность формантной структуры по сравнению с простыми и сильными.

6. А. Долгие В. Средние С. Короткие

6. А. Большая длительность или всей артикуляции (для непрерывных) или же информационно наиболее нагруженной фазы артикуляции (для прерванных); по сравнению с простыми; ярко выраженная формантная структура.

6. В. Средняя длительность артикуляции; средняя отчетливость формантной структуры.

6. С. Краткость артикуляции (только для непрерывных) по сравнению со средними и долгими; бедная и нечетко выраженная формантная структура.

II. ПРИЗНАКИ СПЕКТРАЛЬНОЙ СТРУКТУРЫ.

7. А. Широкодиапазонные В. Среднедиапазонные С. Узкодиапазонные

7. А. Рассредоточение звуковой энергии по широкому диапазону частот спектра с разной плотностью заполнения.

7. В. Концентрация звуковой энергии в диапазонах низких и средних частот.

7. С. Концентрация звуковой энергии в диапазонах низких или высоких частот спектра.

В класс 7. А. из гласных попадают *и* и *е*; из согласных — звонкие турбулентные щелинные переднего образования и аффрикаты такого же образования. В этот класс войдут также глухие турбулентные заднего и глубоководного образования (*x, γ, h, ho*). В класс 7. В. из гласных войдут *a, o*, а из

согласных — сонорные и звонкие смычно-взрывные. В класс же 7. С. из гласных попадает *y*, а из согласных — глухие турбулентные щелинные переднего образования.

8. А. Низкие В. Средние С. Высокие

8. А. Концентрация звуковой энергии в узком диапазоне низких частот; большой объем резонаторов.

8. В. Распределение звуковой энергии в диапазонах низких и средних частот спектра.

8. С. Концентрация звуковой энергии в диапазоне высоких частот спектра. Класс 8. А. объединяет звонкие смычно-взрывные согласные; из гласных туда попадает *y*. В класс 8. В. войдут гласные *a*, и *o*, а из согласных — сонорные. В класс 8. С. попадают глухие щелинные и аффрикаты (условно туда же можно включить и *u*).

9. А. Бемольные В. Простые С. Диззные

9. А. Сдвиг в сторону низких частот всей формантной структуры или преимущественно низких формант; увеличение объема резонатора главным образом за счет огубления.

9. В. Среднее расположение формантов по сравнению с бемольными, с одной стороны, и с дизнными, с другой.

9. С. Сдвиг в сторону высоких частот высоких формант спектра, уменьшение объема резонаторов, расположенных в передней части полости рта.

В класс 9. А. попадают все лабиализованные звуки, а также фарингализованные гласные переднего ряда образования. В класс 9. В. — звуки нелабиализованные, нефарингализованные и непалатализованные; в класс же 9. С. — палатализованные гласные и согласные, а также фарингализованные гласные заднего ряда.

10. А. Нерезкие В. Смешанные С. Резкие

10. А. Наличие хорошо выраженной формантной структуры без участков усиления, недифференцированных по интенсивности.

10. В. Одновременное наличие в диапазоне высоких частот аморфного участка усиления, а в диапазоне низких частот — хорошо выраженной формантной структуры.

10. С. Наличие в диапазоне высоких частот или по всему спектру недифференцированной или очень слабодифференцированной области усиления.

В класс 10. А. войдут все гласные, сонорные и звонкие смычно-взрывные фонемы. В класс 10. В. — звонкие щелинные и аффрикаты. В класс 10. С. — все глухие щелинные как переднего, так и заднего образования.

III. ПРИЗНАКИ АКУСТИЧЕСКИХ ГРАДАЦИЙ

Согласные фонемы разных локальных рядов в зависимости от величины и конфигурации резонаторов, расположенных впереди места артикуляционного сужения или смычки, можно сгруппировать по следующим акустическим градациям:

1. Б-подобные (билабиальные смычно-взрывные и щелинные, а также сонорный *m*).
2. Ф-подобные (дентолабиальные щелинные и аффрикаты).
3. Д-подобные (дентальные смычно-взрывные, а также сонорный *n*).
4. С-подобные (денто альвеолярные щелинные и аффрикаты).
5. Ш-подобные (альвеолярные щелинные и аффрикаты).
6. Й-подобные (среднетвердонебные щелинные).
7. Л-подобные (латеральный сонорный *l* и система латеральных щелинных и аффрикат).
8. Г-подобные (постпалатальные resp. превелярные смычные и щелинные).
9. Х-подобные (поствелярные щелинные и верхнефарингальные аффрикаты).
10. h-подобные (среднефарингальные и ларингальный щелинные).

DISCUSSION

Wenk:

Во второй графе ваших дифференциальных признаков некоторые термины повторяются („смешанные“, „простые“, „средние“). Не может ли это мешать однозначности описания фонемы?

Wiede:

В очень интересном докладе предлагается вместо бинарной классификации звуков трехчленная. Такая классификация дает возможность исключить подгруппы бинарных делений. Но, как докладчик сам сказал, некоторые из установленных им дифференциальных признаков занимают промежуточное положение. Можно ли такие признаки считать дифференциальными? Существуют ли в конкретных языках такие оппозиции как „назальный“ — „назализованный“; „звонкий“ — „придыхательный“? Нельзя ли сказать, что на практике трехчленная классификация все-таки сведется в бинарную.

Skoumal:

„Градация“, которая соответствует делению фонем по месту образования, выносится за рамки акустической классификации, предложенной в докладе. Правильно ли это?

Гаприндашвили:

Акустические градации по существу соответствуют локальным рядам. Разницу между ними составляет объединение в одну градацию, порою, больше одного локального ряда. Поэтому нами использована акустическая терминология.

Трехчленное противопоставление: звонкий — придыхательный — смычногортанный, ясно, что нельзя обнаружить в языках, где смычные и щелинные группируются в парные системы. В указанном противопоставлении объединены два бинарные противопоставления теории дифференциальных признаков: звонкий — глухой и глоттализированный — неглоттализированный.

Ответ на вопрос Скоумала: Акустические градации у нас по существу соответствуют т. н. локальным рядам (несмотря на то, что некоторые из них объединяют по два и больше локальных рядов). Здесь вопрос только терминологии.

Трехчленные противопоставления и акустические градации представляют собою главные и вспомогательные координаты, соответственно, для выделения отдельных фонем из классов фонем.

Ответ на вопрос Венка: Повторение терминов: „смешанные“, „простые“, „средние“ не означает повторение самых дифференциальных признаков в разных трехчленных оппозициях. Например, термин „смешанные“ в трехчленном противопоставлении: голосовые — смешанные — шумные обозначает класс фонем, характеризующихся наличием, одновременно, и голоса и шума. Этот же термин в оппозиции: нерезкие — смешанные — резкие используется для обозначения класса фонем с аморфной областью усиления в диапазоне высоких частот спектра и рельефно дифференцированной областью усиления в диапазоне низких частот. Термин „простые“ в оппозициях: сильные — простые — слабые, долгие — простые — краткие и бемольные — простые — диззные используется для обозначения класса исходных единиц при противопоставлений фонем по долготе и краткости, в первом случае, по напряженности и слабости, во втором, и по признаку наличия сдвига частот в спектре вверх или вниз, в третьем.

Ответ на вопрос Э. Виде: В бинарной классификации дифференциальных признаков нет места, например, для назализованных (т. е. орально-назальных), которые в ряде языков (нахские языки иберийско-кавказской группы языков и др.) представляют собою самостоятельные фонемы. Бинарная оппозиция: непрерывный — прерванный игнорирует класс пульсирующих щелинных (вибрантов *γ, x, p*), а оппозиции: звонкий — глухой и неглоттализированный-глоттализированный по существу представляют собою одну трехчленную оппозицию по признаку участия голосовых связок в артикуляции: звонкий — придыхательный — непродувный (смычногортанный).

Каждый член трехчленной оппозиции, будучи фонологически самостоятельным, регулярно противопоставляется остальным двум. Нельзя говорить о „промежуточном положении“ ни одного из звеньев трехчленных оппозиций, так как каждое звено представляет самостоятельный класс фонем. Следовательно, исключается возможность сведения трехчленных оппозиций к бинарным.

THE DIAGNOSTIC FUNCTION OF IDIOSYNCRATIC PHONETICS

PAUL L. GARVIN*

This paper is programmatic, in the sense that it discusses proposed research, rather than reporting on research already completed.

By idiosyncratic phonetics, I mean the nondistinctive variations in the pronunciation pattern of individual speakers within the frame of reference of the phonological pattern of the speech community to which they belong. These phonetic idiosyncrasies are not considered random fluctuations, but are assumed to have nonlinguistic psychological correlates; it is because of these correlates that I claim that idiosyncratic phonetics has a diagnostic function.

I find this assumption operationally significant, since it permits me to consider the use of idiosyncratic phonetic features as diagnostic tools—more specifically, I claim that these features constitute, or at least approximate, actual diagnostic parameters which can be used to measure, or to contribute to the measurement of, psychologically significant characteristics of the speakers.

The broader conception on which this assumption is based is the view that language constitutes a "projective system," in the sense that the use of language by the individual can be as revealing of personality characteristics as his responses to specifically designed projective tests. In the present paper, I shall restrict myself to the phonological aspects of this projective function of language.

From the operational standpoint that I am taking, the key questions are the following two: (1) which are the particular measurable phonological features that can be expected to serve as diagnostic parameters, and how; (2) what research design is best suited to test the validity of my assumptions?

Phonological features can serve as diagnostic parameters in two senses. Either, particular features within the same constant (or near-constant) speech pattern have diagnostic value, or the speech pattern as a whole has diagnostic value, as is the case when the speaker engages in code-switching. Both of these aspects of the projective function of speech may, of course, be manifested jointly in the same speech situation.

Diagnostic parameters within the same speech pattern include the following:

Rate of speech, that is, the number of phonological units per unit of time.

* State University of New York at Buffalo, U.S.A. This paper was written while the author was at the Bunker—Ramo Corporation, Canoga Park, California, U.S.A.

Frequency and length of pauses.

Hesitation phenomena, including false starts.

Frequency and spacing of different intonation patterns.

Pitch span, that is, the difference in pitch level between the highest and the lowest pitch used in a given intonation pattern; the same intonation pattern may be uttered with pitch spans of different width.

Code-switching as a diagnostic parameter involves primarily a change from formal to informal and regional pronunciation patterns.

The parameters presented above are listed in the order of their tractability, that is, their susceptibility to measurements and other techniques of analysis.

The most tractable parameter is rate of speech. It can be established by the simple tabulation of phonological units (phonemes, syllables, breath groups, depending on the investigator's convenience and interest), or their orthographic representation (such as typewritten transcript with indication of elapsed time). Inaccuracies in the delimitation of the phonological units can be expected to cancel each other out if the sample is large enough.

Rate of speech is probably the only parameter that can be measured without taking into account the details of prosodic patterning. In the measurement of the remaining parameters, the prosodic pattern of the language has to be taken into account to an increasing extent.

The frequency of pauses can likewise be established by simple tabulation, but in this case, the orthographic representation is a good deal less reliable, since the orthography (in those languages where it does so) usually indicates those pauses that are desirable from a normative standpoint, and not those that are actually present in a stretch of speech. The length of pauses, of course, cannot be established by simple tabulation but requires the measurement of the actual phonetic event. In order to be maximally meaningful, the tabulation of pause frequencies and the measurement of pause length should be related to their interpretation in terms of phonemic junctures, although I contend that even without such an interpretation these diagnostic parameters should be revealing if they can be correlated with psychological states.

Hesitation phenomena again can best be observed and tabulated in their phonetic manifestation, unless their orthographic representation in a transcript is unusually exhaustive. In view of the wide-spread habit of subconscious editing of written matter, this will be the case only of the transcripts prepared by specially-trained personnel who have been made aware of the significance of these phenomena for the interpretation of the psychological state of the speaker. Another difficulty is the well-known fact that hesitation phenomena are not uniform, and hence different types of these phenomena will have to be tabulated, and evaluated, separately. This not only requires a prior awareness of the prosodic pattern and the cultural pattern of responses to speech, in order to determine the presence of hesitation phenomena, but it also presupposes some prior psycholinguistic assessment of the

types of hesitation phenomena that can be expected in the speech pattern under investigation.

The frequency and spacing of intonation patterns can obviously be determined only on the basis of a preceding detailed prosodic analysis. Both the frequency and the spacing of intonation patterns must be measured, not merely in terms of units of time elapsed, but also in terms of the number and length of appropriate phonological units such as phonemes, syllables, breath groups, over which any given pattern extends. This also presupposes that the type of phonological unit with which a given type of intonation pattern is associated be ascertained in the course of the prior prosodic analysis.

The pitch span has to be ascertained on the basis of a prior knowledge of the intonation pattern that exhibits a given span. This presupposes not only a prosodic analysis leading to a categorization of intonation patterns into distinctive types, but also the establishment of the norms in terms of which the pitch span is to be measured and interpreted.

In the case of code-switching, the very concept implies that the speech pattern as a whole, or rather, the speaker's change from one speech pattern to another, has a diagnostic function. The situation is complicated by the well-known fact that code-switching need not be complete—shorter or longer stretches, or just individual features, of one pattern may be embedded in the other just as readily as there may be a complete transition from speech using one pattern to speech using another. The extent to which switching takes place is in itself of diagnostic value, as well as the projective significance of the actual patterns.

From an operational standpoint, it is not necessary to observe or measure all the characteristics of the patterns involved in code-switching. What is important is to determine which phonological features are sufficiently characteristic to serve as reliable indices of the fact that code-switching has taken place, and of the extent to which it has taken place.

As was noted, the diagnostic parameters discussed above have been placed along a tentative tractability scale. The operational significance of this scale lies in the fact that it is quite conceivable that the grosser, more tractable parameters (such as rate of speech and frequency and length of pauses) may have a diagnostic value equal to the more subtle and less tractable ones.

On the basis of some preliminary study, the following general research design is suggested for the investigation of the diagnostic function of idiosyncratic phonetics:

1. At least in the early stages of the research, individual voices should be studied separately.
2. For each voice to be studied, a sample of spontaneous speech should be obtained under controlled conditions in an emotionally neutral environment and at a time when the speaker is known to be in a psychological state as close to emotionally relaxed as possible. The parameters of these samples will serve as reference points for the study of samples obtained under other conditions.

3. Samples of spontaneous speech of the same voices should be obtained under equally controlled conditions, but in an environment which is not emotionally neutral, and/or at a time when the speaker is not in an emotionally relaxed condition.

4. Samples of the two types are to be compared in terms of all the parameters listed in order to determine the differences in diagnostic value between various parameters.

5. If the different parameters, as is conceivable, turn out to have approximately equal diagnostic value, then only the gross parameters should be used in further studies.

6. If the different parameters turn out to have significantly different diagnostic value, then obviously the most revealing parameters should be used in further studies.

DISCUSSION

Hill:

I should like to ask how closely the parameters described correspond to H. L. Smith's paralinguistic signals.

W. R. Lee asked whether Prof. Garvin would consider adding to his list of parameters syllable-, word-, or possibly phrase-length, or also loudness and softness of the voice.

Garvin:

Ad Hill: The problem of paralinguistics is a theoretical one. My paper is concerned with the operational problem of tractability, irrespective of the theoretical interpretation of the particular feature under consideration.

Ad Lee: Syllable length certainly constitutes a valid parameter. Words and phrases, on the other hand, do not constitute phonological, but rather grammatical units, and are therefore not included in this discussion.

RHETORISCHE ANALYTIK

HELLMUT GEISSNER

Anzahl und Bedeutung der in Akten rhetorischer Kommunikation erzeugten gesprochenen Sprachwerke motivieren die Notwendigkeit rhetorischer Analyse. Ältere Untersuchungen hatten nur den schriftlichen Wortlaut als Quelle, bestenfalls ein Stenogramm. Mit dem kritischen Instrumentarium der ‚literarischen Rhetorik‘ konnte Rede nur als Literatur untersucht werden. Die Reduktion auf Inhalte entsprach der allgemeinen Verkennung des Gesprochenen und genügte den Interessen bestimmter Historiker, Politologen und Linguisten.

Die technische Möglichkeit, Reden als gesprochene auf Tonträgern zu speichern, schaffte neue Quellen. Diese neuen Quellen verlangen neue Methoden der Analyse. Ein Schritt auf diese Hermeneutik des Gesprochenen war Bühlers Modell der einen triadischen semantischen Relation. Mit diesem, der ‚Rhetorik‘ des Aristoteles entlehnten, Modell ließen sich Kundgabe-, Darstellungs- und Appellfunktion und ihr jeweiliges Verhältnis verdeutlichen. Entscheidender Bezugspunkt der Analyse blieb aber die Intention des Redners. Dies erklärt, warum die vorwiegend am Sprechausdruck interessierten Psychologen, jüngst auch Biophonetiker, sich hier orientierten.

Einen Schritt über diesen Ansatz hinaus führt die Verwendung des im Anschluß an Peirce entwickelten ebenfalls triadischen semiotischen Funktionsmodells. Der ‚semantische‘ Richtungsstrahl behält die Funktion Zeichen: Bezeichnetes, der ‚pragmatische‘ die Funktion Zeichen: Zeichenverwender; der dritte — der ‚syntaktische‘ — kennzeichnet die materiale Funktion Zeichen: Zeichen, die am alten Modell nicht darstellbar war. Die akustische Quelle dokumentiert am ‚Großzeichen‘ Rede dreierlei: Sprachliches, Sprecherisches und Rhetorisches (s. str.). Die rhetorische Analytik muß folglich in jedem der drei semiotischen Parameter jeweils Sprachliches, Sprecherisches und Rhetorisches teils quantifizieren, teils qualifizieren und ihre Korrelationen bestimmen als Konstituierende des Großzeichens Rede.

Sprachliches und Sprecherisches beeinflussen sich wechselseitig intensiv. So bestimmt z. B. die Silbigkeit der Wörter das Sprechtempo oder die Anzahl finiter Verben die Pausierung; umgekehrt löst eine bestimmte Art der Akzentuation Langsätze in kürzere Denkschritte auf und steigert deren Wirksamkeit, oder die inverse Intonation macht die Aussage zur Frage. Es ist nicht nur der Sprechstil vom Sprachstil abhängig, so daß es genügen könnte, diesen zu analysieren und jenen als sekundäre Modifikation abzutun, sondern gerade bei der Rede, zumal bei der frei in der Situation

Tabelle 1

1. syntaktisch	2. semantisch	3. pragmatisch
1.1 sprachlich	2.1 sprachlich	3.1 sprachlich
1.1.1 Wortanzahl, Silbigkeit/Wort, m Silbenzahl, Silben-Entropie	2.1.1 begriff. Inventar, Terminologie	3.1.1 Verständlichkeit der Wörter und Sätze (vgl. 3.3.2)
1.1.2 Wortmaterial: Fachsprachen, Fremdwörter, Wortbildungen	2.1.2 Worthäufigkeit: Verhältnis innerhalb der Rede zur allgem.	3.1.2 Wortwirksamkeit, Eindrucksstelle, Ausdrucksstelle, Wiederholung
1.1.3 Sprachstufen: Hochsprache, Umgangssprachen, Mundart(en)	2.1.3 funktionaler Aspekt, bevorzugte Wortbildungen, Stellenwert der Wörter, Schlüsselwörter, Anredepronomen usf.	3.1.3 Satzwirksamkeit, Wiederholung, Verkürzung (vgl. 2.1.4)
1.1.4 Satzanzahl, Wörter je Satz, m Satzlänge	2.1.4 Stil: Denkfunktion verbal—nominal, parataktisch, hypotaktisch	3.1.4 Satzarten: Funktion z. B. dialogisch, befehlend, fragend, werbend usw.
1.1.5 Satzarten, Setzungen, Einfachsatz, Satzgefüge	2.1.5 Satzarten in ihrem Aussagewert: Urteil, Frage, Wunsch, Behauptung	3.1.5 Satzarten in ihrer Intention, z. B. rationalisierend, emotionalisierend
1.2 sprecherisch	2.2 sprecherisch	3.2 sprecherisch
1.2.1 Lautung: Sprechstufen (hochgelaftet, umgangssprachlich, mundartlich) Deutlichkeit	2.2.1 sinnkonstituierende Funktionen der 1.2-Elemente, z. B. Pausierung, Kadenzierung, Tonhöhenbewegung	3.2.1 Lautung: situative Varianten der Sprechstufen (Hörerschaft)
1.2.2 konstitutionelle Audsruksqualitäten: Stimmlage, -klang, -fülle, -fehler Lautungsfehler	2.2.2 Bedeutungsfunktion von 1.2.2, vor allem Tonerhöhung, Klangfarben, z. B. Ironie, Brutalität, Güte; Divergenz/Konvergenz von 2.1 und 2.2	3.2.2 Hörbarkeit (Raumart, -größe, -akustik)
1.2.3 Verlaufsqualitäten: Tempo, -wechsel, Pausierung, Akzentuierung, Tonhöhenbewegung	2.2.3 Intensitätsgrade, Außenspannung, Innenspannung, Intentionalität (Sache, Selbst, Hörer)	3.2.3 situative und rollenspezifische Varianten von 1.2.2 u. 1.2.3
1.3 rhetorisch	2.3 rhetorisch	3.2.4 Aktualisieren der 2.1 und 2.2 Stimmungen
1.3.1 Redeaufbau Gliederung	2.3.1 Redeaufbau: Funktion der Glieder	3.2.5 Hörer-Reaktionen: verbale, nicht-verbale
1.3.2 Umfang der Glieder	2.3.2 Funktion der rhetorischen Figuren	3.2.6 Reaktionen auf die Hörer-Reaktionen
1.3.3 Rhetorische Figuren	2.3.3 Argumentationsfolge	3.3 rhetorisch
1.3.4 Argumente (Zahl)	2.3.4 Redehaltung: monologisch, dialogisch (fiktiv, virtuell, aktuell)	3.3.1 Redeaufbau: Wirkung
1.3.5 Gesamtumfang (-dauer)	2.3.5 Redeart, (-ziel)	3.3.2 Wirksamkeit der Argumentationsfolge
		3.3.3 Hörerbezug
		3.3.4 Redesituation: Ort, Zeit, Hörerschaft
		3.3.5 Rededauer: Wirkung (hörbare Handlungsimpulse der Hörerschaft usw.)

sprechdenkend entwickelten, ist der Sprachstil vom Sprechstil abhängig. Beide sind beeinflusst vom Redestil, den sie beeinflussen, so daß auch *das* ‚Rhetorische‘ nach den drei Hinsichten zu analysieren ist.

Die rhetorische Analytik berücksichtigt demnach in den drei Zeichenfunktionen (syntaktisch, semantisch, pragmatisch) jeweils Sprachliches, Sprecherisches und Rhetorisches. Sie kann auf diese Weise Rede qua Rede einem adaequaten hermeneutischen Prozeß zuführen, ohne vorschnell auf die Psychologie des Redners, die z. B. politische Situation, die Soziologie der Hörerschaft oder heteronome Wertsysteme auszuweichen. Sie kann außerdem signifikante Unterschiede zwischen verschiedenen Reden eines Redners, zwischen verschiedenen Rednern und verschiedenen Epochen bestimmen. Weiterhin liefert sie der Redetheorie Daten für die Definition und Klassifikation der Redearten. Schließlich ist sie anwendbar bei der Analyse von als Kurzreden verstandenen Gesprächsbeiträgen im Verlauf formalisierter Gespräche, jedoch nicht zur Analyse ganzer Gesprächsverläufe; dazu sind noch andere Kriterien erforderlich.

DISCUSSION

Kurka:

Die ausgeführte Analytik als systematischer Entwurf kann m. E. nur als Voraussetzung für die wissenschaftliche Bewertung einer Rede gelten, die den gesellschaftlichen Bezug einbeziehen muß. Zu den angeführten Faktoren erhebt sich die Frage, wie „Innen“- und „Außenspannung“ eines Redners wissenschaftlich erfaßbar sind.

Schnorrenberg:

Angesichts der vom Ref. aufgezeigten und wohl ohne Zweifel zu Recht bestehenden Forderung nach größerer Einbeziehung des bisher wohl überwiegend als sekundär betrachteten Sprechstils: in eine Analyse des Gesprochenen, erscheint die gleichfalls erhobene Warnung vor einem Ausweichen in die Psychologie des Sprechers, hier des Redners, sowie in die Soziologie der Zuhörerschaft, nicht einsehbar, da m. E. ersteres notwendigerweise die zuletzt genannten Bezüge postuliert.

Geißner:

Ich trug keine experimentell gewonnenen Ergebnisse vor, sondern einen systematischen Entwurf. Danach muß aus dem Katalog der Faktoren und aus der Faktorenanalyse die mögliche Wertung zunächst ausgeklammert werden; ebenso die Psychologie des Redners und Soziologie der Hörer, obwohl sie z. T. in 2.2 und 3.2, 3.3 vorkommen. Mein Entwurf und die begonnene Arbeit versucht ohne derartige Rückgriffe „Rede als Rede“ zu analysieren.

ÉTUDE EXPÉRIMENTALE DES ASPIRÉES SONORES EN MALAYALAM

JEAN-DENIS GENDRON*

I. OBJET ET MÉTHODE

Meillet dit des aspirées sonores de l'indo-européen, qu'elles „se distinguaient des sonores simples, sans doute par une articulation glottale qu'on n'a pas le moyen de déterminer exactement.“¹ La présente communication, qui a pour objet l'étude du fonctionnement des cordes vocales dans la réalisation des aspirées sonores en malayalam, voudrait éclairer quelque peu cette question.

Le malayalam, langue du groupe dravidien parlée au sud de l'Inde se prête bien à une telle étude: il possède en effet vingt occlusives, différenciées entre elles par le lieu d'articulation, la présence ou l'absence de sonorité, ou encore d'aspiration, soit:³

	Labiales	Dentales	Préalatales	Cacuminales	Vélaires	
non-aspirées	<i>p</i>	<i>t</i>	- <i>č</i>	<i>ʈ</i>	<i>k</i>	sourdes
	<i>b</i>	<i>d</i>	<i>ǰ</i>	<i>ɖ</i>	<i>g</i>	sonores
aspirées	<i>pʰ</i>	<i>tʰ</i>	- <i>čʰ</i>	<i>ʈʰ</i>	<i>kʰ</i>	sourdes
	<i>bʰ</i>	<i>dʰ</i>	<i>ǰʰ</i>	<i>ɖʰ</i>	<i>gʰ</i>	sonores

Pour l'expérience, nous avons comparé entre elles les articulations figurant dans ce tableau, en les plaçant, soit en position initiale de mot, soit en position intervocalique. Le sujet⁴ les a prononcées dans des mots isolés, à voix

* Université Laval (Québec, Canada).

¹ Meillet, Antoine, *Introduction à l'Étude Comparative des Langues indoeuropéennes*, Paris, Hachette, 1953, p. 88.

² Meillet A. et Cohen, M., *Les Langues du Monde*, Paris, Champion, 1952, pp. 487—501.

³ Voir G. Straka, *Album Phonétique*, Québec, Les Presses de l'Université Laval, planche 51, pour l'appareillement de ce système avec celui du sanscrit.

⁴ Le Père Gabinus Petta, âgé d'environ cinquante ans, né et élevé dans le comté d'Ernakulam, au centre de l'Etat de Kerala, de parents originaires de la même circonscription. Le malayalam est la langue maternelle de chacun d'eux. Le Père Petta vit actuellement à Thevara, comté d'Ernakulam, Etat de Kerala.

un peu plus forte que la moyenne, en laissant un silence entre chacun d'eux.⁵

La méthode oscillographique convenant bien au but que nous poursuivions, nous avons fait des tracés à quatre lignes, *Lm* (larynx-microphone), *Bm* (bouche-microphone), *Ba* (bouche-air) et *Nm* (nez-microphone),⁶ qui ont été mis sur pellicules et lus à l'aide d'un lecteur de microfilms.

II. RÉSULTATS

a) Comparaison des aspirées sonores et sourdes en position initiale de mot.

En position initiale de mot, les aspirées sonores se présentent en malayalam, comme partiellement sonores, mais avec prédominance de la sonorité sur la sourdité, au cours de l'occlusion. L'explosion et l'aspiration sont dévoisées, comme pour les aspirées sourdes (voir figure 1).

La mesure de la partie sourde et sonore de l'occlusion fait voir que cette dernière est en moyenne deux fois plus longue que la première; cette proportion peut varier de huit fois à une fois dans les réalisations particulières, avec, parfois, des durées égales pour chacune des parties, ou même, mais rarement, une portion sonore plus brève que la sourde. Ce sont des consonnes demi-sonores.

Cette portion sonore de la tenue, peu audible au premier abord pour une oreille étrangère, paraît fonder essentiellement l'opposition aspirées sonores et aspirées sourdes, en malayalam, chez le sujet dont nous avons étudié l'articulation.

b) Comparaison des aspirées sonores et sourdes en position intervocalique.

1. Tendance à réaliser les consonnes comme sonores à l'intervocalique

Le système d'oppositions ci-dessus décrit, semble fonctionner de la même façon, lorsque les aspirées sont à l'intervocalique, et ce, malgré une tendance manifeste à réaliser les consonnes comme sonores dans cette position.

En effet, sur dix-sept réalisations, à l'intervocalique, de consonnes dites sourdes non aspirées, neuf ont été complètement sonorisées, les huit autres l'étant partiellement, à des degrés divers. (Voir figure 2.)

On pourrait s'attendre que la tendance agisse de la même façon sur les aspirées sourdes, et, à fortiori, sur les aspirées sonores: elle se manifeste, mais à un degré beaucoup moindre que pour les non aspirées sourdes.

⁵ En tout, vingt réalisations pour les aspirées sonores, dix-huit, pour les aspirées sourdes, dix-huit, pour les non-aspirées sonores et trente-quatre pour les non-aspirées sourdes, dont vingt pour la consonne *p*; les réalisations se partageant à peu près également entre les positions initiale et intervocalique.

⁶ Voir G. Straka, op. cit., p. 162, pour la description de ce dispositif technique, les microphones étant, pour nos expériences, reliés à un oscillographe à galvanomètres (de marque *Consolidated Electrodynamics Corporation*) au lieu d'inscripteurs électro-mécaniques. L'air phonatoire (*Ba*) a été inscrit sur les tracés à l'aide d'un dispositif mis au point par M. Lucien Lebourhis, technicien du laboratoire de phonétique expérimentale de l'Université Laval.

2. Comportement des cordes vocales pendant la tenue des aspirées sourdes et sonores, à l'intervocalique

Les aspirées sourdes. Elles se sonorisent, mais peu, la partie sourde de l'occlusion étant en moyenne 4.4 fois plus longue que la partie sonore. Ce sont des sourdes légèrement sonorisées (Voir figure 3).

Les aspirées sonores. Loin de se sonoriser totalement, elles se comportent presque comme les aspirées sourdes: elles ont une tenue sourde généralement un peu plus longue que leur tenue sonore, parfois égale ou un peu plus brève. (Voir figure 3). La différence tient à ce que la partie sonore de l'occlusion est en moyenne presque trois fois plus longue que pour les aspirées sourdes, de telle sorte que la proportion entre les deux parties de l'occlusion est réduite à 1.4 au lieu de 4.4.

Pour les aspirées sonores, s'il n'y a pas, en durée, prédominance de la sonorité pendant l'occlusion, cette sonorité paraît être plus fermement et plus régulièrement marquée que pour les aspirées sourdes.

Ceci suffit-il à fonder l'opposition entre aspirées sonores et sourdes à l'intervocalique? Il semble que oui, pour le sujet dont nous avons étudié l'articulation.

c) Fonctionnement des cordes vocales pour les aspirées sonores en malayalam.

D'après Rousselot,⁷ il existe des aspirées complètement sonores, c'est-à-dire, avec occlusion et explosion sonores, pour lesquelles la grande dépense d'air faite au moment de l'aspiration ne gêne pas le mouvement vibratoire des cordes vocales. Tel n'est pas le cas chez notre sujet.

Tirant argument du fait que, dans cette langue, contrairement à la tendance qui se manifeste pour les non-aspirées sourdes, les aspirées sourdes et même sonores ne se sonorisent que partiellement à l'intervocalique, on peut penser que la préparation de la glotte pour l'aspiration inhibe le mouvement vibratoire des cordes vocales, provoquant un arrêt de celui-ci avant la rupture de l'occlusion buccale.

Or, le mouvement vibratoire des cordes vocales peut être empêché de deux façons: par occlusion glottale ou par écartement des cordes vocales sous l'effet, ou encore, en prévision d'une grande dépense d'air. On sait en effet que cette dernière est, dans la parole, en partie fonction de l'ouverture de la glotte.⁸ Il est par ailleurs généralement admis, dans le cas des aspirées sourdes, que les cordes vocales sont écartées pendant la tenue de ces consonnes,⁹ bien qu'on ait trouvé, dans certains parlars, qu'elles pouvaient s'articuler avec une glotte fermée pendant l'occlusion.¹⁰ Qu'en est-il en malaya-

⁷ Rousselot, *Principes de Phonétique expérimentale*, Tome II, pp. 886—891.

⁸ Voir G. Straka, *Respiration et Phonation*, dans *Bulletin de la Faculté des Lettres de Strasbourg*, mai—juin 1957, pp. 401—404, pour un excellent exposé sur les variations de débit et de dépense d'air dans la chaîne parlée et leurs causes.

⁹ Voir en particulier J. Chlumsky, dans *Revue de Phonétique*, T. III, pp. 396—399 et dans *Časopis pro moderní filologii*, 1925, pp. 31—33 et 1927, pp. 120—121.

¹⁰ Voir A. Schmitt, in *Teuthonista*, T. VII, 1931, pp. 287 sq. et F. Falc'hun, *Le système consonantique du breton*, 1951.

lam, pour les aspirées sonores? La glotte est-elle fermée ou ouverte, entre la fin des vibrations laryngales et la rupture de l'occlusion buccale?

La lecture de la ligne laryngale sur les tracés oscillographiques conduit à des interprétations diverses:

1. dans certaines réalisations des aspirées sonores, en particulier pour les *t'* et *d'*, apparaît sur la ligne laryngale, au moment de la rupture de l'occlusion buccale, une déviation abrupte, qu'on pourrait interpréter comme une explosion laryngale, résultant d'une brusque réouverture de la glotte, après occlusion ferme de celle-ci (voir figure 4); mais, d'une part, cette supposée explosion laryngale n'apparaît qu'irrégulièrement, et, d'autre part, des expériences faites en laboratoire font voir qu'elle peut provenir de la répercussion de l'explosion buccale au niveau du larynx;

2. dans d'autres réalisations, celles des *k'* et des *g'* la rupture de l'occlusion buccale est suivie, au niveau du larynx, d'une ou deux déviations, apparentées par leur forme à l'onde sonore d'un ton laryngien, et qu'on pourrait interpréter comme étant produites par de faibles battements des cordes vocales; mais ces battements ne fournissent pas davantage de renseignements sur la position des cordes vocales avant la rupture de l'occlusion buccale, car ils pourraient aussi bien se produire la glotte étant ouverte¹¹ qu'après brusque réouverture de celle-ci¹² (Voir figure 4.);

3. dans la réalisation des autres aspirées sonores, on ne voit généralement sur la ligne du larynx que des déviations faibles et désordonnées, résultant du frottement de l'air entre les cordes vocales pendant l'aspiration (voir figure 5). Donc rien ici encore qui permette de déterminer la position des cordes vocales avant la rupture de l'occlusion buccale.

Des phénomènes signalés ci-dessus et de l'interprétation qu'on en a tentée, il est pratiquement impossible de déterminer avec sûreté la position des cordes vocales, entre la fin du mouvement vibratoire de celles-ci et la rupture de l'occlusion buccale. c'est-à-dire, au cours de la dernière portion de la tenue des aspirées sonores en malayalam.

III. CONCLUSION

Tout au plus pourrait-on penser, en se fondant cette fois-ci sur le rapport connu entre le fonctionnement des cordes vocales et la dépense d'air¹³, que l'arrêt précoce du mouvement vibratoire de ces dernières dans l'articulation des aspirées sonores en malayalam est fait en prévision de la grande dépense d'air que constitue l'aspiration;

¹¹ Voir dans le film *Vocal Cord Action in Speech*, réalisé par Elizabeth T. Uldall, du département de phonétique de l'Université d'Edinburgh, le comportement des cordes vocales au moment du passage du *h* sourd à la voyelle qui suit.

¹² Voir dans le même film le comportement des cordes vocales après le coup de glotte.

¹³ Voir G. Straka, *Respiration et Phonation*, op. cit. pp. 401—404.

et supposer alors que les cordes vocales, pour mieux réaliser celle-ci, restent écartées, dans la position du *h* sourd,¹⁴ une fois les vibrations terminées.

Ce même rapport permettrait d'expliquer de façon satisfaisante, outre la demi-sonorité des aspirées sonores en malayalam, les transformations dont les aspirées sonores de l'indo-européen ont été l'objet: a) conservation de la sonorité et perte de l'aspiration en iranien, slave, baltique, albanais et celtique; b) conservation de l'aspiration et perte de la sonorité en grec.¹⁵

DISCUSSION

Fischer—Jørgensen:

J'ai commencé une investigation instrumentale des consonnes occlusives de différentes langues indiennes, et j'ai observé une variabilité assez grande des aspirées sonores. On peut trouver ou occlusion sonore + aspiration sonore ou occlusion sonore + aspiration sourde ou bien (moins souvent) occlusion sourde + aspiration sonore, et aussi des cas intermédiaires.

¹⁴ Voir dans le film déjà cité la position des cordes vocales pour l'articulation du *h* sourd; de même G. Straka, *Album Phonétique*, op. cit., planche 5.

¹⁵ Voir Meillet, op. cit. p. 88.

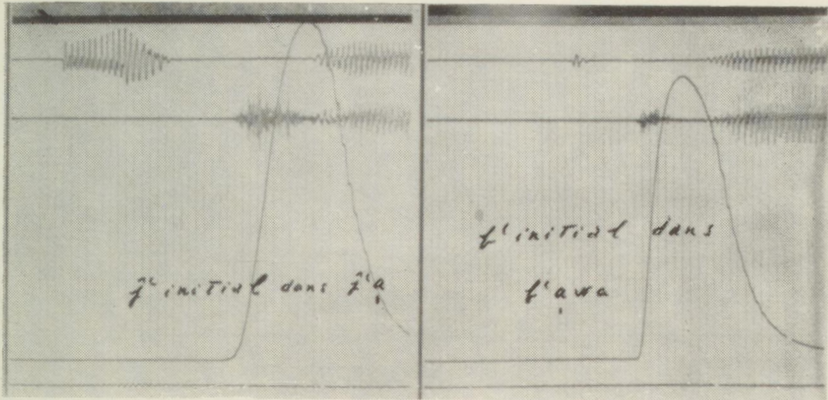


Fig. 1.

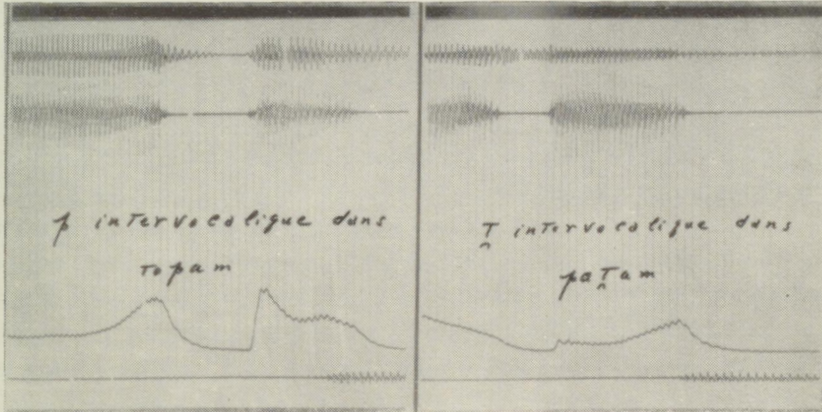


Fig. 2.

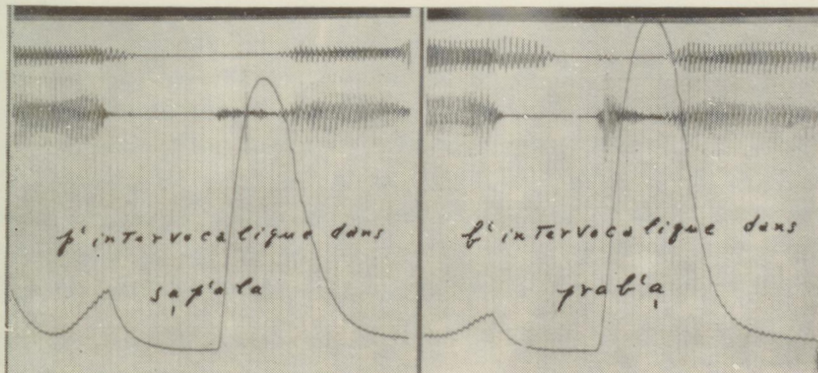


Fig. 3.

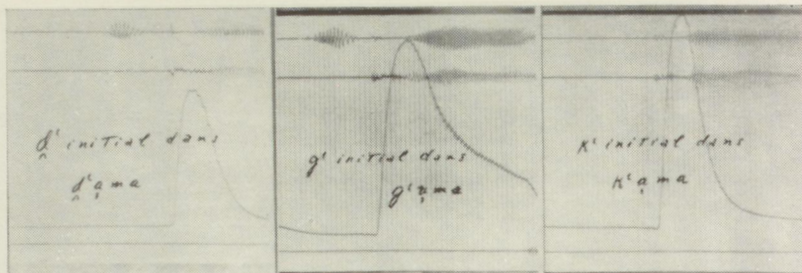


Fig. 4.

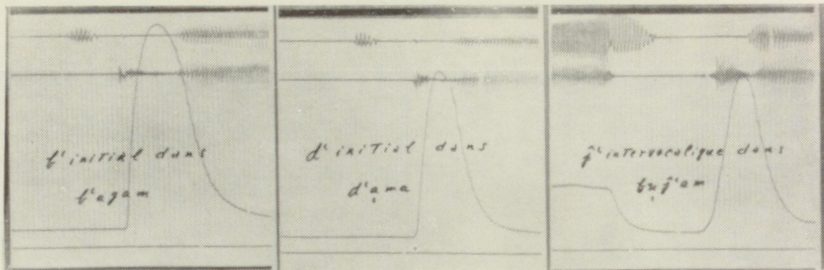


Fig. 5.

PHONEME COGNITION AND PERCEPTION

DON GEORGE*

I used to think that in any language the phonemes were distinguished from each other on the basis of the physical characteristics of the sound spectrum and were identified by selected features which made each phoneme distinctive. Therefore, any deviation from the permitted allophonic variations in the language would shift the phoneme boundary and produce a significantly different utterance. Consequently, I devised numerous minimal-pair drills and loaded sentences in an attempt to condition the auditory and articulatory mechanisms of students to distinguish accurately those sound complexes to be cognized as the same from those to be cognized as different.

Meanwhile, I was a little bothered by the question of how we know when a person makes a phoneme substitution, or a non-permissible mutilation, which phoneme was intended. This ability, I argued, came as a result of phonetic training, and the argument became the rationale for courses in phonetics.

Some time ago in an advanced phonetics class, after I had been discussing some of the predictable phonetic problems of both native and non-native speakers and the importance of accurate auditory and articulatory drills to correct these "errors", an astute student asked, "Why is it that when a non-native speaker shifts phonemes or mutilates a phoneme even people with no phonetic training usually know what he is trying to say?" In other words, how does the phonetically untrained person cognize signals which are physically different as being the same, signals which are physically the same as being different, cognize missing elements in the signal and ignore excrescent elements?

After admitting the statistical necessity of controlling variables in experimental situations, and considering the various para-linguistic clues to meaning and the possible significance of the interaction of uncontrolled variables in normal communication, we attempted to arrive at an understandable answer to the question. The thesis which evolved may be stated somewhat as follows: Phoneme cognition is as much, if not more, dependent upon a complex of extra-phonetic clues as upon the physical characteristics of the acoustic signal. These clues may operate independently

* University of Southern Mississippi.

or together, sequentially or simultaneously, and serve to validate normal cognition and to correct faulty acoustic signals.

If we take as a model for phoneme perception a series of monitors which regulate and classify the components of the total acoustic input, first censoring, or filtering out, any unwanted elements, and then, for the remainder, identifying the frequencies physically present, measuring the relative strength of the frequencies, and establishing the temporal relationships of the components, we have the basis for identifying the phonetic structure of an utterance whether it is cognized or not. This model, though, is inadequate to describe cognition. The precision with which the identification can be made is, of course, a matter of phonetic sophistication. Normal cognition—that is, speakers using the same code system—involves scanning the memory for lexicon, syntactical and morphological pattern, collocatability of items, situational probability, and adding to all these the evidence from other senses.

In spite of the large number of “emes” that have been coined since the term *phoneme* was introduced, I would like to add another to refer to the minimal element necessary to transmute a mutilated or missing acoustic signal into a cognizable unit. I would suggest the term *cogneme*.

As a simple illustration let me use a sentence in which at least three categories of cognitive correction are needed. Suppose one hears, “I cut a peak feesh.” One program consisting of the total lexical inventory scans the utterance and reports, “*Feesh* not in lexicon; [iy] = [i]. All lexical items now stand. But a second program of permitted collocations declares, “*Peak* not collocatable with *fish*.” Another scanning of the lexicon locates no probable compounds with *fish*, the stress pattern program indicates a qualifier, an inventory of possible qualifiers gives the minimal correction as, [p] = [b], [iy] = [i], and [k] = [g]. Evidence from other senses confirms the validity of *big fish*.

If a signal from the visual area reports a large, bloody knife, then *cut* is cognized as valid. If, on the other hand, a large, cleaned, wrapped fish is visually perceived the correction may be made, “Evidence for *cut* negative. Possession indicated. Change [k] to [g] and [ʌ] to [a].” Thus *cut* is cognized as *got*. But, if a freshly caught fish and fishing tackle is visually perceived, the report may come in as, “Evidence for *cut* negative. Possession negative. Change [ʌ] to [ɔ].” Thus, *cut* is cognized as *caught*.

Since this scanning is done almost instantaneously and performed sub-consciously, we are not aware of having done it, and our cognitive system hears the statement, “I caught a big fish.”

In addition to cognemes of lexicon, collocation, and perceived context already mentioned, there are those which deal in the same way with permitted versus not-permitted syntactical sequences and morphological structures of the type, “Me go now,” or “He give me it.” We cognize *me* as *I*, supply a missing auxiliary, shift a present to a past morpheme, and re-arrange the sequence of direct and indirect objects.

Though the perceived referent or the observed context may be all that is needed when the situation makes the referent obvious (as when a Bulgarian friend of mine said, “It’s nize here on your *teh-rahss*” when we were sitting on my terrace), when there is no visible referent or context, cognition will depend on a knowledge of possible versus not-possible referents. If more than one referent is possible, a tentative correction on the basis of probability will be made and a second cogneme will be elicited to confirm or negate the decision. If, for example, in a non-contextual situation you are asked, “You like docks?” either *dogs* or *ducks* may be considered equally probably, and the reply will likely be evasive until a subsequent cogneme establishes the referent more clearly.

Finally, such clues as vocal quality, facial expression, gesture, or some supra-segmental phoneme may serve as cognemes by revealing the intention of the speaker. Thus, when a Louisiana Cajun calls out, “Hey, you goddam dog?” even if we do not know he is searching for his lost hounds we know from one of these clues that he is not cursing us but is asking if we have his dogs.

In summary, though there is nothing profoundly new in these concepts which have been discussed by many writers, I have found that this model, a system of monitors processing the raw data of the physical stimulus, accepting, rejecting, substituting, adding, or taking from, according to the various levels of the total linguistic experience is a helpful device for explaining to students the complex way by which we cognize incomplete or inaccurate signals. This corrective ability permits a considerable degree of linguistic tolerance. Without it communication would be difficult, if not impossible, unless absolute phonemic conformity were achieved.

IST DIE THESE „LAUTWANDEL DURCH INNERE KAUSALITÄT“ SPRACHGEOGRAPHISCH BEWEISBAR?

JAN GOOSSENS

Die ersten Versuche, mit dialektgeographischen Argumenten nachzuweisen, daß Sprachwandel sich durch innere Kausalität vollziehen könne, hat Gilliéron gemacht. Das bekannteste Beispiel ist der Aufsatz über die Bezeichnungen für die Katze und den Hahn in der Gaskogne, der im Buch *Etudes de géographie linguistique* von Jules Gilliéron und Mario Roques 1912 erschien. Die Beweisführung besteht hier in der Darlegung der Koinzidenz zweier Isoglossen, welche die Verbreitung zweier sprachlicher Erscheinungen abgrenzen, zwischen denen ein kausaler Zusammenhang vermutet wird. Die erste ist das Lautgesetz, das Homonymie zwischen den Fortsetzungen von lat. *cattus* und lat. *gallus* in der Gaskogne hätte verursachen müssen. Die zweite ist das Auftreten von anderen Wörtern als die lautgesetzliche Fortsetzung von *gallus* zur Bezeichnung des Hahns in einem Teil des betreffenden Gebietes. Aus der Tatsache, daß die Isoglosse des Lautgesetzes und die der Ersatzbezeichnungen für *gallus* (wenigstens teilweise) zusammenfallen, wird auf einen kausalen Zusammenhang zwischen beiden Erscheinungen geschlossen: *gallus* sei als Folge der Homonymenfurcht durch andere Bezeichnungen ersetzt worden.

Die Darlegung des Zusammenfalls von Isoglossen hat Moulton als Beweismethode übernommen in seinem Aufsatz *Lautwandel durch innere Kausalität: die ostschweizerische Vokalspaltung* (ZfMaf 28 (1961), S. 227—251). Er hat dabei versucht, die Beweiskraft des Verfahrens zu steigern, indem er nicht zwei, sondern drei Erscheinungen mit koinzidierenden Grenzen in die Untersuchung einbezogen und den Verlauf einer Linie aus dem zweier anderer zu erklären versucht hat. Im folgenden möchte ich das Wesentlichste aus seinen Ausführungen zusammenfassen.

Das mhd. Kurzvokalsystem hatte nach der Durchführung des Umlauts die folgende Gestalt:

<i>i</i>	<i>ü</i>	<i>u</i>
<i>e</i>	<i>ö</i>	<i>o</i>
<i>ë</i>		()
<i>ä</i>		<i>a</i>

Dieses System war asymmetrisch: die Vorderreihe enthielt vier Stufen, die Hinterreihe nur drei. „Die Asymmetrie konnte auf zwei Weisen behoben werden: entweder durch die Schaffung eines neuen Hintervokals, wodurch das leere Fach gefüllt wurde;

oder aber durch den Zusammenfall zweier Vordervokale, wodurch das leere Fach einfach verschwand“ (S. 241).

Das mhd. Langvokalsystem vor der Dehnung der alten Kürzen war symmetrisch. Es sah folgendermaßen aus:

\bar{i}	$\bar{ü}$	\bar{u}
\bar{e}	\bar{o}	\bar{o}
$\bar{æ}$		$\bar{ā}$

Die Tatsache, daß die Zahl der Öffnungsgrade im Kurzvokalsystem in der Vorder- und in der Hinterreihe nicht die gleiche war, brachte also eine doppelte Unregelmäßigkeit mit sich: Erstens eine Asymmetrie zwischen Vorder- und Hinterreihe im Kurzvokalsystem; zweitens eine Asymmetrie zwischen Kurz- und Langvokalsystem.

Moulton zeigt, daß in den ostschweizerischen Mundarten das Gleichgewicht überall nachträglich wiederhergestellt wurde. Das geschah jedoch nicht immer auf die gleiche Weise. Die einfachste Lösung war der Zusammenfall zweier Vordervokale im kurzen System, wodurch die vordere wie die hintere Reihe drei Stufen enthielt. Diese Lösung kam in den zwei folgenden Varianten vor:

A $i \rightarrow [i]$	B $i \rightarrow [i]$
$e \rightarrow [e]$	$e \rightarrow [e]$
$\bar{e} \rightarrow [æ]$	$\bar{e} \rightarrow [æ]$
$\bar{ü} \rightarrow [æ]$	$\bar{ü} \rightarrow [æ]$

Es entstand hier jeweils ein Vierecksystem mit drei Öffnungsgraden, das mit dem Langvokalsystem vollkommen harmoniert:

i	$\bar{ü}$	u
e	\bar{o}	o
$\bar{æ}$		$\bar{ā}$

Dieses System findet sich im Westen des Untersuchungsgebiets. Im Osten (C) dagegen hat sich bei den Kurzvokalen der Vorderreihe nichts geändert; im Norden (D) wurde das alte $\bar{ä}$ gehoben, bis es mit \bar{e} zusammenfiel. Das Ergebnis dieser letzten Entwicklung mußte folgendes asymmetrisches Dreiecksystem sein:

i	$\bar{ü}$	u
e	\bar{o}	o
ε	()	
a		

Aber gerade in den Gebieten mit den Entwicklungen C und D setzte sich die „ostschweizerische Vokalspaltung“ durch, das heißt eine Senkung bei den Kurzvokalen, die folgendermaßen dargestellt werden kann:

$i \rightarrow [i]$	$\bar{ü} \rightarrow [\bar{ü}]$	$u \rightarrow [u]$
$\rightarrow [e]$	$\rightarrow [o]$	$\rightarrow [o]$
	$\bar{o} \rightarrow [o]$	$\bar{o} \rightarrow [o]$
	$\rightarrow [e]$	$\rightarrow [e]$

Die Senkung war anfänglich allophonischer Art, aber durch die Wirkung der Analogie und durch spätere Entwicklung postvokalischer Konsonante wurde die Opposition $[o] \neq [e]$ (und $|\bar{o}| \neq |\bar{e}|$) phonologisiert.

Die Vokalspaltung findet sich überall, wo die Entwicklung der Vorderreihe keine Symmetrie im System herbeigeführt hatte (m. a. W., in den Gebieten C und D). Andererseits kommt sie nirgends in den Gebieten vor, wo durch die Entwicklung der Vorderreihe eine Symmetrie im System entstanden war (m. a. W., in den Gebieten A und B). In der Tatsache, daß die Abgrenzung der Vokalspaltung nahezu vollkommen mit jener der Entwicklungen C und D zusammenfällt, erblickt Moulton einen Beweis für seine These, daß die Spaltung sich durch innere Kausalität, das heißt durch Systemzwang vollzogen hat.

Es kommt hinzu, daß sich im Gebiet der Vokalspaltung ein vierstufiges Langvokalsystem findet, und in den Bereichen ohne Spaltung ein dreistufiges. Die Zahl der mhd. Längen konnte sich durch die Dehnung alter Kürzen vermehren, wenn nämlich die neuen Längen nicht mit den alten zusammenfielen. In diesem Zusammenhang scheint vor allem die Entwicklung des gedehnten a von Bedeutung zu sein. Dort, wo es mit der alten Länge $\bar{ā}$ zusammenfiel (X), enthielt die hintere Reihe des Langvokalsystems drei Stufen, wie im Mhd. Wo es die Stelle von $\bar{ā}$ einnahm, das seinerseits in velarer Richtung auswich und mit \bar{o} zusammenfiel (Y), enthielt die hintere Reihe ebenfalls drei Stufen. Es kommt aber in einem Gebiet, das gerade von der Spaltung bei den Kurzvokalen überlagert wird, noch eine dritte Entwicklung vor (Z): Bei der Dehnung von a wich das alte $\bar{ā}$ in der Richtung von \bar{o} aus, blieb jedoch halbwegs zwischen \bar{o} und $\bar{ā}$ stehen. In diesem Gebiet enthält das System vier Hintervokale: $\bar{ü}, \bar{o}, \bar{e}, \bar{ā}$.

Durch die ostschweizerische Vokalspaltung wurde also eine doppelte Symmetrie im Vokalsystem wiederhergestellt: 1. Die zwischen Hinter- und Vorderreihe bei den Kürzen; 2. Die zwischen dem System der Kürzen und jenem der Längen.¹ Angeblich wird die innere Kausalität der Spaltung durch eine doppelte Koinzidenz von Isophonen bewiesen: 1. Die Grenzen des Spaltungsbereiches fallen mit denen des Gebietes zusammen, wo bei der Entwicklung der Vorderreihe der Kurzvokale keine Symmetrie zustande gekommen war. 2. Sie fallen ebenfalls zusammen mit denen des Bezirks, wo ein vierstufiges Langvokalsystem entstanden war.

¹ Moulton sagt allerdings nicht, ob der Parallelismus so weit geht, daß in Gebiet C das vierstufige Langvokalsystem viereckig und in Gebiet D dreieckig ist.

Moulton hat jedoch offenbar seine Beweisführung nicht für vollkommen befriedigend gehalten. In einem Zwischensatz auf S. 247 schreibt er: „(insofern solche Dinge überhaupt „bewiesen“ werden können)“. Seine Schlußfolgerung scheint mir auch nicht zwingend zu sein. Sie wäre es gewesen, wenn er hätte zeigen können, daß die Vokalspaltung sich in ihrem ganzen Verbreitungsgebiet zu gleicher Zeit vollzogen hätte. Aber diese Hypothese ist selbstverständlich unbeweisbar. Aus der Koinzidenz der Isoglossen läßt sich zwar schließen, daß die *Verbreitung* der Spaltung irgendwie von sprachinternen Faktoren abhängig ist, aber nicht ohne weiteres, daß diese Lautentwicklung durch innere Kausalität *entstanden* ist. Die Verbreitung der betreffenden Erscheinung kann ja noch auf eine andere Weise gedeutet werden. Es ist nämlich denkbar, daß sie zu einer bestimmten Zeit an einem bestimmten Punkt innerhalb des heutigen Spaltungsbereiches entstanden ist. Ob sie sich hier durch innere Kausalität oder infolge anderer, extern-linguistischer Faktoren vollzog, ist angesichts der Beweisführung Moultons irrelevant. Dieses Zentrum hätte dann die Neuerung über die Umgebung ausgestrahlt. Die Spaltung wäre von den umliegenden Orten übernommen worden und hätte dort überall eine doppelte Symmetrie im Vokalsystem wiederhergestellt. Die Verbreitung der Spaltung durch Strahlung hätte dann schließlich die Grenze eines Gebietes erreicht, wo sie nicht nur die Symmetrie im Vokalsystem nicht wiederhergestellt, sondern vielmehr eine auf andere Weise zustandegekommene doppelte Symmetrie zerstört hätte. An dieser Grenze wäre die Neuerung dann stehen geblieben. Jenseits der Schranke hätte das Vokalsystem durch die Übernahme der Spaltung ja schlechter funktioniert.

Welche der beiden Interpretationen die richtige ist, kann die Sprachgeographie nicht beweisen. Wenn es richtig ist, daß die Dialektgeographie in den letzten Jahren im Untermauern von Hypothesen der linguistischen Theorie eine neue Aufgabe gefunden hat, so scheint mir Moultons Aufsatz zur Unterstützung einer solchen Annahme kein wesentlich neues Element beigetragen zu haben. Beide denkbare Deutungen des Kartenbildes enthalten ein gemeinsames Element, dessen Realität für gesichert gehalten werden kann: Verschiedene Lautentwicklungen können sich gegenseitig verhindern. Wo die Vokalspaltung auftrat, war das Entstehen bzw. das Fortbestehen von Dreierreihen in den Kurz- und Langvokalsystemen ausgeschlossen, und umgekehrt, wo Dreierreihen zustandegekommen waren, konnte die Spaltung sich nicht durchsetzen. Eine negativ definierte Teleologie in der Lautentwicklung, also die Annahme einer den Systemen innewohnenden Tendenz, nicht „schlechter“ zu funktionieren, scheint mir also durchaus vertretbar.

Zu diesem Schluß war die Dialektgeographie jedoch schon früher gekommen. Schon 1951 hatte Weijnen (*Taalgeografie en interne taalkunde*. Amsterdam 1951, S. 21) darauf hingewiesen, daß die Grenzen der Diphthongierung von germ. *i* und *û* und der Entwicklung von *ê* zu *î* und von *ô* zu *û* in Limburg weithin zusammenfallen, und daß jenseits der gemeinsamen Grenzlinie *î* und *û* sowohl als *ê* und *ô* erhalten blieben oder wiederhergestellt wurden. „Man könnte also sagen: Im Gebiet, wo aus altgerm. *eo* (sic) und *ô* ein *ê* bzw. *ô* entstand, konnten *î* und *û* nicht diphthongiert werden,

weil sonst die extremen Stellungen im Phonemsystem unbesetzt wären. Oder aber: das brabantische Sprachsystem hat die durch Diphthongierung verschwundenen alten *î*- und *û*-Laute durch das *u* und *i* von *Buch* und *Brief* ersetzt.“ Das Verdienst der Arbeit von Moulton scheint mir vor allem darin zu liegen, daß sie ein von Weijnen nur andeutungsweise skizziertes Verfahren unter Berücksichtigung des fast vollständigen Vokalsystems der untersuchten Dialekte und mit vielen Besonderheiten erläutert, in musterhafter Weise angewendet hat. Daß eine Sprache ihr Phonemsystem durch innere Kausalität ändern kann, hat er jedoch nicht bewiesen.

FORMANTBEWEGUNGEN IN VOKALEN DEUTSCHER MUNDARTEN DES WESTLICHEN HARZES

JOACHIM GÖSCHEL*

Der folgende Beitrag steht im Rahmen der Bemühungen, die in der Abteilung Phonetik des Deutschen Sprachatlasses unternommen werden, um die Artikulationsbasis in den deutschen Mundarten zu ergründen. Ziel dieser gemeinsamen Bemühungen ist, Kriterien für eine exakte Bestimmung der Artikulationsbasis zu finden, die sich letzten Endes auch kartographisch fixieren lassen. Daraus erklärt sich auch die Wahl des geographischen Gebietes in der vorliegenden Untersuchung. Im westlichen Harz treffen nämlich drei verschiedene Mundartgruppen zusammen (Abb. 1). Differenzierungsmerkmale müßten also auch auf phonetischer Ebene relativ leicht sichtbar gemacht werden können.

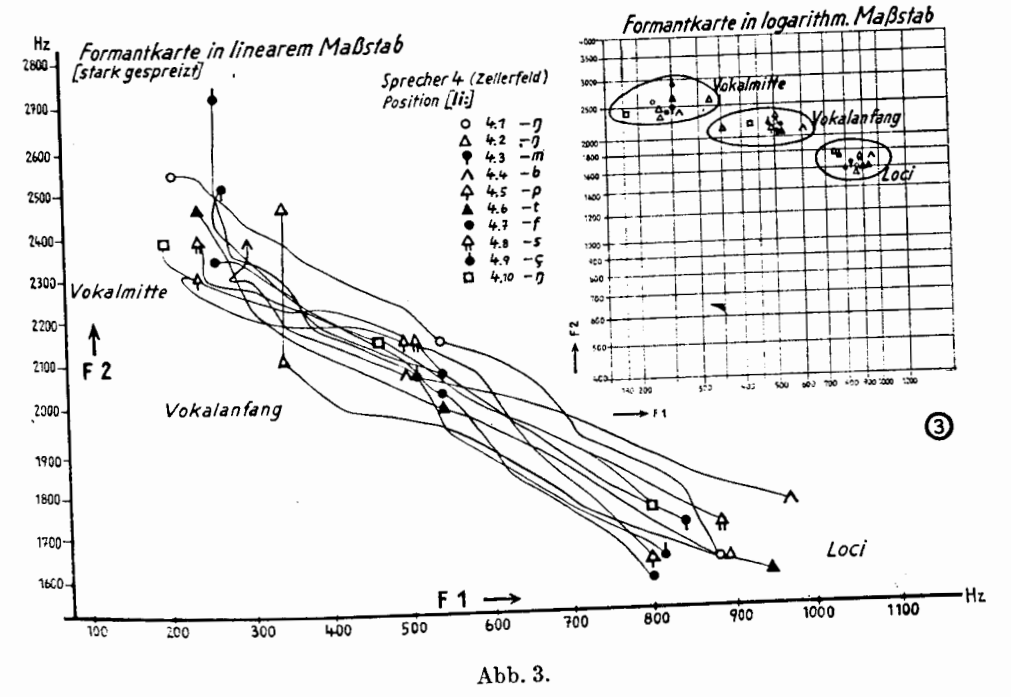
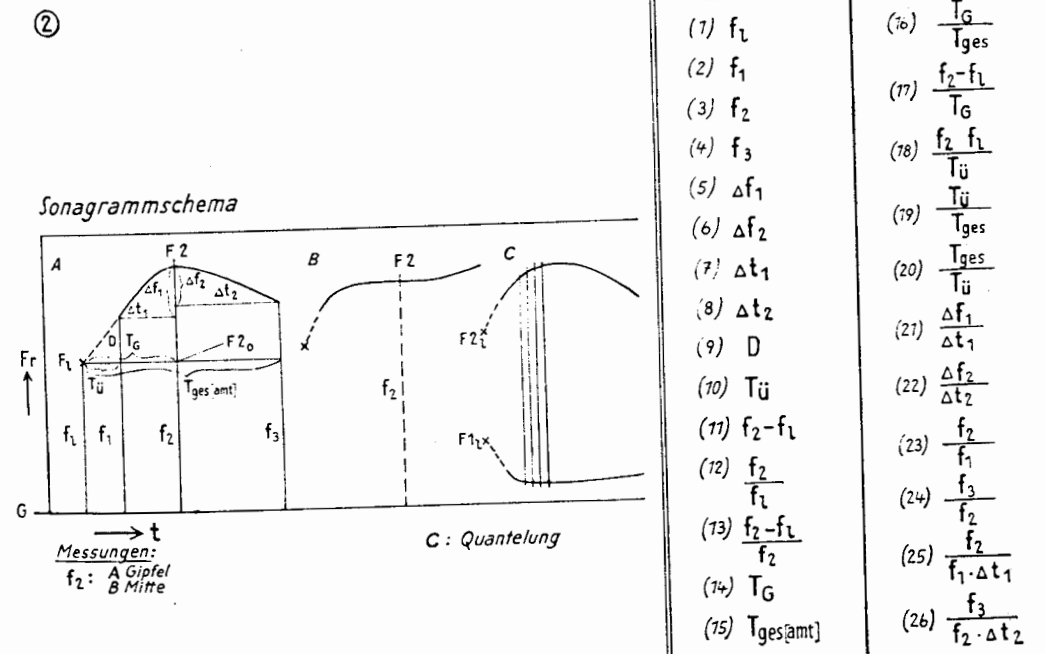
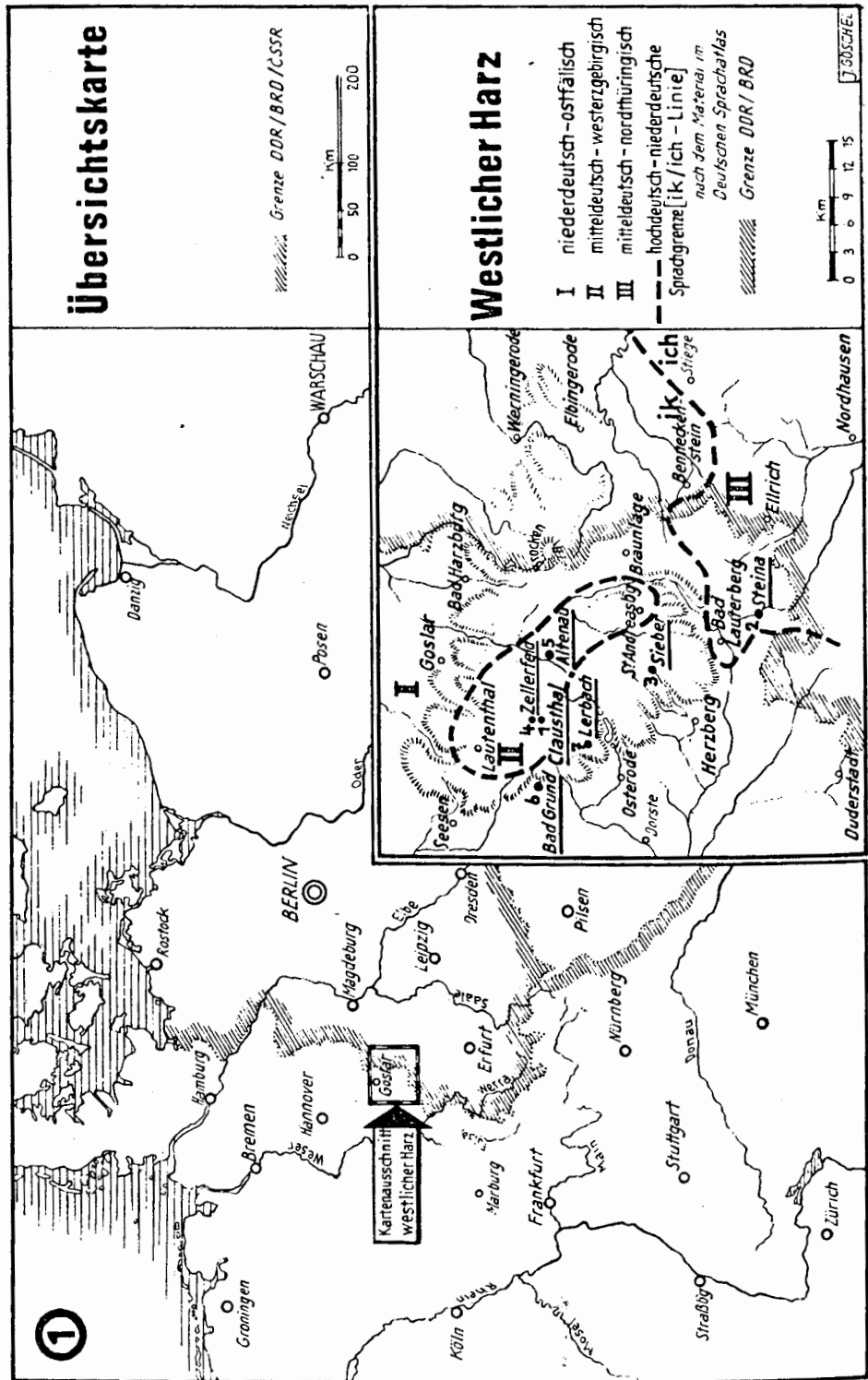
Es soll hervorgehoben werden, daß die vorliegenden Erörterungen tastende Versuche und methodische Voruntersuchungen darstellen. — Ausgangspunkt waren Sonagramme mit auffällig gekrümmten Formantverläufen. Diese Bewegungen im Hinblick auf allgemeine Gesetzmäßigkeiten und eventuelle Unterscheidungsmerkmale innerhalb der Mundartgruppen zu untersuchen, war Ziel und Absicht der Versuche.

Arbeitsmethodisch wurde folgender Weg eingeschlagen. Als gemeinsame Ausgangsbasis für sieben männliche Sprecher aus sieben verschiedenen Ortsmundarten (Abb. 1) wurde die Lautgruppe [li: + Kons.] gewählt. Das Material wurde aus technisch hochwertigen Tonbandaufnahmen — aus Einzelwortabfragungen und fortlaufender Erzählung — entnommen.¹ Die Anzahl der verfügbaren Beispiele pro Sprecher war unterschiedlich (erste Zahl = Sprecher, zweite Zahl = Beispiele): 1. 8, 2. 7, 3. 5, 4. 10, 5. 11, 6. 10, 7. 3. Von diesen Beispielen wurden Sonagramme hergestellt,² von denen die Verläufe von F1 und F2 mit den Loci des [1-] auf trans-

* Aus der Abteilung Phonetik des Forschungsinstituts für deutsche Sprache 'Deutscher Sprachatlas', Universität Marburg an der Lahn. Dem Abteilungsleiter G. Heike danke ich für zahlreiche Aussprachen zu den im Text behandelten Fragen.

¹ Es wurden Aufnahmewagen, Aufnahmegeräte und Schallträger des Deutschen Spracharchivs, Münster/Westf., benutzt. Die technischen Daten können entnommen werden aus E. Zwirner, Anleitungen zu sprachwissenschaftlichen Tonbandaufnahmen (= Lautbibliothek der deutschen Mundarten 31). Göttingen 1964. S. 43—48.

² Verwendet wurde ein Sonagraph vom Typ Vibralyzer (Vibration Analyzer) der Fa. Kay Electric, Pine Brook, N. J./USA und Sonagrammpapier Type B. Die Filtereinstellung war „weit“ (190 Hertz).



parentes Millimeterpapier übertragen wurden. An diesen Bewegungen wurden die verschiedensten Messungen vorgenommen und in Beziehung zueinander gebracht (Abb. 2). Für jeden Meßwert wurde eine Rangfolge unter den Sprechern aufgestellt. — Außerdem sollten die Bewegungen, die F1 und F2 innerhalb eines Vokals unter Einschluß der Loci durchlaufen, genau verfolgt werden. Für diesen Zweck wurden die Formantbewegungen gequantelt und die Werte für F1 und F2 alle zwei Millimeter abgelesen. Die Werte wurden auf linear weit gespreizte Formantkarten eingetragen (Abb. 3). — Solche Diagramme wurden für alle Sprecher und eine Versuchsperson G (mit mitteldeutsch-ostthüringisch gefärbter Hochsprache), die die Wortbeispiele reproduzierte, hergestellt.

Die Auswertung brachte noch nicht die erhofften Resultate von Unterscheidungsmerkmalen der drei Dialektgruppen. — Aus der Vielzahl der Meßergebnisse ist eins von allgemein-phonetischem Interesse. Im Koordinatensystem, wenn auf der Ordinate T_G und auf der Abszisse T_{ges} aufgetragen sei, liegen die Mittelwerte der Sprecher auf einer ansteigenden Geraden, während die Einzelwerte etwas streuen. Eine Abweichung zeigt lediglich Sprecher 1, bei dem vermutlich ein Meßfehler vorliegt, der aber trotz mehrmaligen Nachprüfens noch nicht gefunden werden

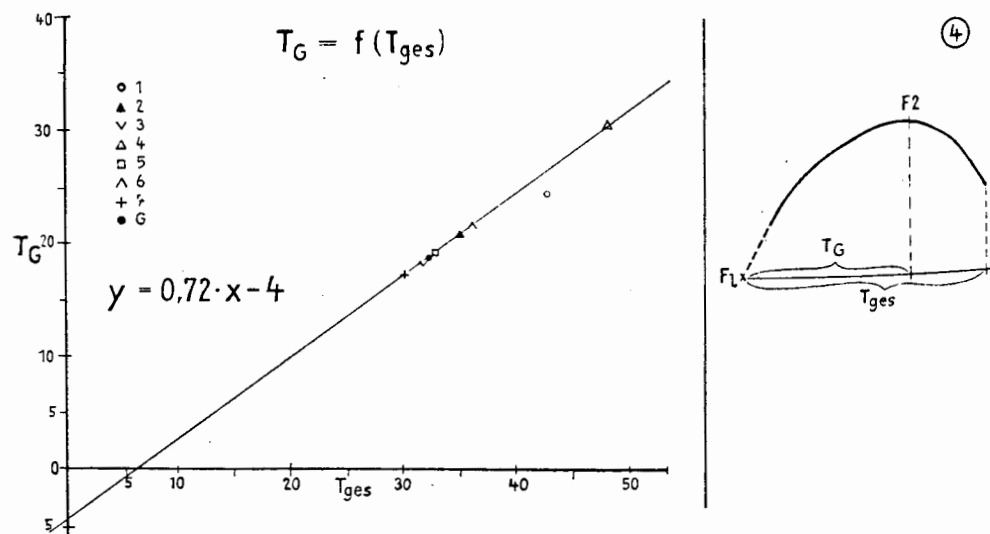


Abb. 4.

konnte (Abb. 4). Dieses Verhältnis zeigt, daß mit wachsendem T_G auch T_{ges} zunimmt. Aus einem gegebenen T_G kann T_{ges} (oder umgekehrt) nach der Formel $y = 0,72 \cdot x - 4$ vorausgesagt werden. Bei verändertem Sprechtempo bleibt demnach T_G in einem bestimmten Verhältnis zu T_{ges} .

Ohne auf die Einzelheiten der internen Meßdaten hier näher eingehen zu können, zeichnen sich insgesamt gesehen in der Überzahl solche Gruppierungen innerhalb der sieben Sprecher ab, die auf eine ziemlich deutliche Absonderung der Sprecher

1 und 4 hinauslaufen. Entweder schließen sich beide zusammen und liegen peripher oder sie bilden polare Extremwerte. Dieses Bild vermittelt allerdings schon die bloße Rangfolge der Meßdaten. Im Hinblick auf eine Differenzierung der drei Mun-

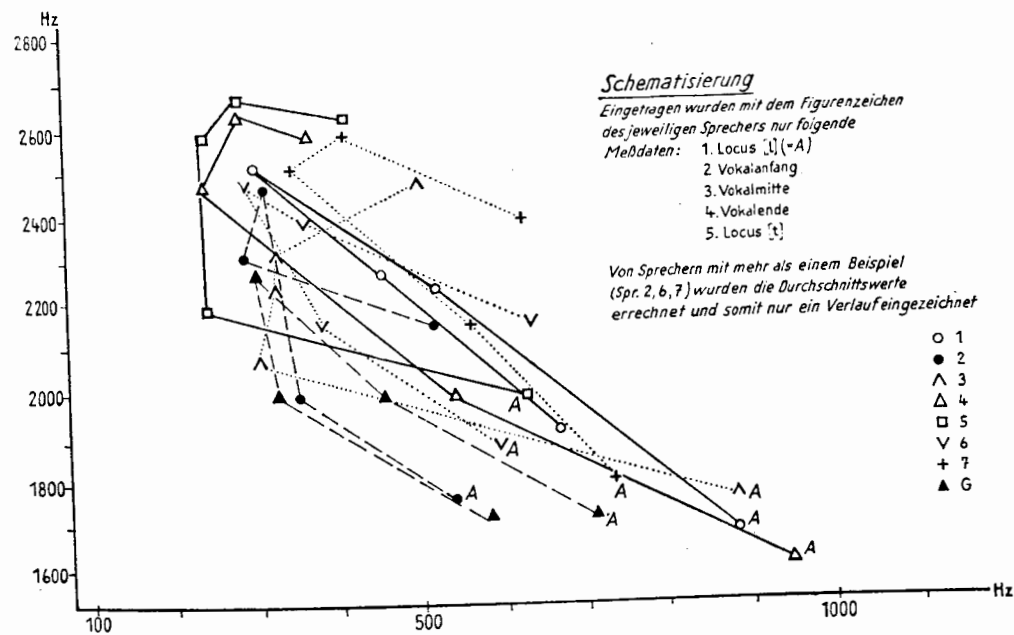
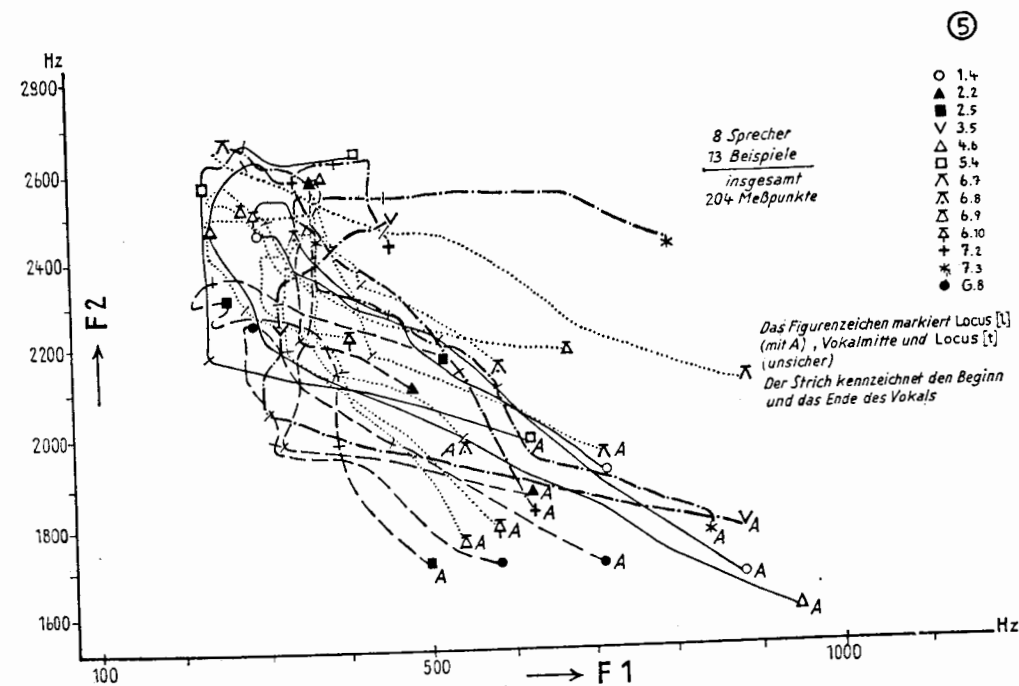


Abb. 5.

arten bedeutet das ein Absondern der beiden mitteldeutsch-westerzgebirgischen Ortsmundarten von Clausthal (1) und Zellerfeld (4). Der dritte Sprecher dieser Gruppe (Altenau [5]) zeigt diese Merkmale jedoch nicht. Auch die drei Sprecher des Niederdeutschen (Sieber [3], Grund [6], Lerbach [7]) bilden hinsichtlich der Meß-

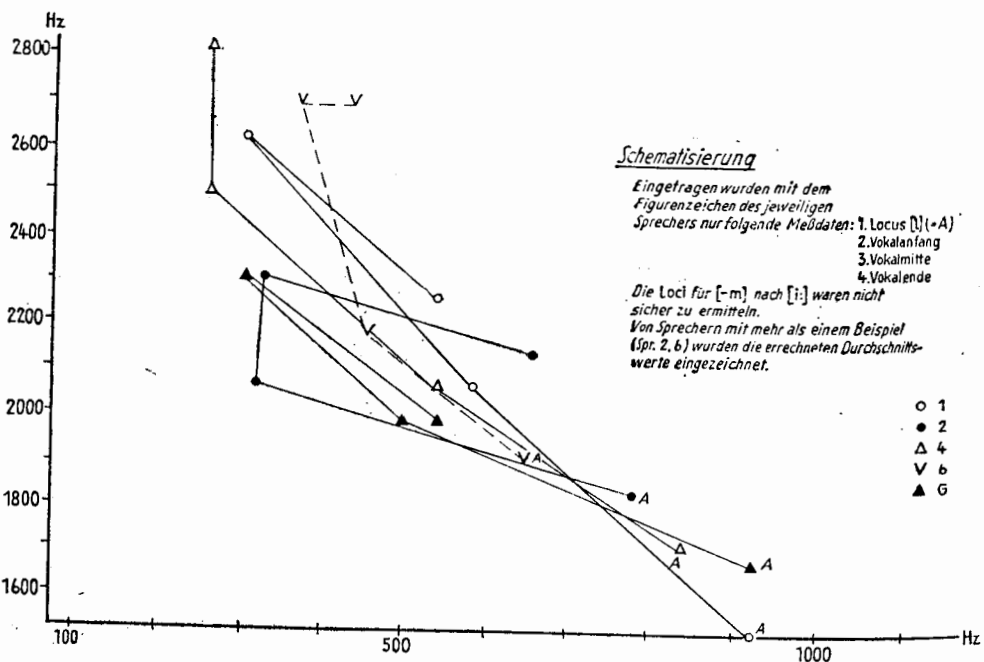
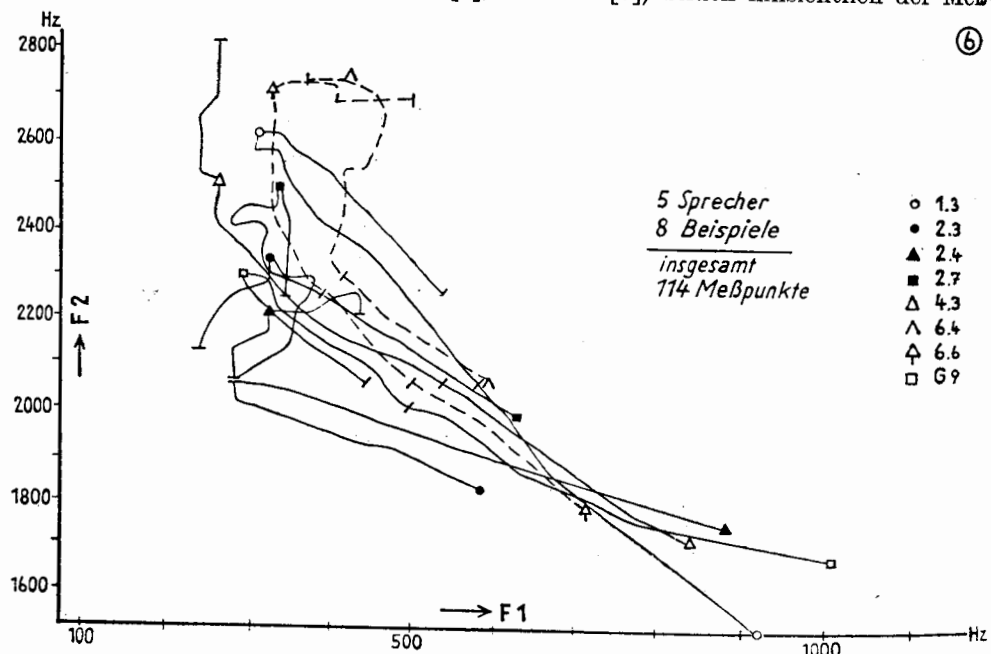


Abb. 6.

werte keine Einheit. Und Sprecher 2, der den mitteldeutsch-nordthüringischen Dialekt (Steina) vertritt, nimmt keine Sonderstellung ein.

Der nächste Schritt, der, vom Material her gesehen, methodisch vielleicht verfrüht war, verfolgte die Bewegungen von F 1 und F 2 auf weit gespreizten Formantkarten bei acht Sprechern im gleichen lautlichen Kontext von [li:t] mit 13 Beispielen und [li:m] bei fünf Sprechern mit acht Beispielen. — Eine Tendenz zur Einheitlichkeit der Bewegungsverläufe beider Formanten war weder innerhalb des gleichen lautlichen Kontextes aller Sprecher noch für mehrere Beispiele eines Sprechers erkennbar (Abb. 5: Sprecher 2, 6, 7; Abb. 6: Sprecher 2 und 6). Die Schematisierung dieser Bewegungen unter Verwendung der Durchschnittswerte von jenen Sprechern, die mehr als ein Beispiel liefern, zeigt im Diagramm darunter zwar ein klareres Bild, aber noch keine charakteristischen Strukturen. Analoge Verläufe können für 4. 6 und 5. 4, sowie für 6. 7 und 7. 3 bei [li:t] registriert werden. — Die im ersten Teil der Untersuchung festgestellte Absonderung der Sprecher 1 und 4 war hier nicht gegeben.

Für die Bewegungen in [li:t] und [li:m] wurden durch Auflegen von Deckblättern folgende Abschnitte des zeitlichen Verlaufs sichtbar gemacht: Loci [l], Vokalanfang, Vokalmitte, Vokalende und bei [li:t] die Loci von [t]. Die Lage der Loci des auslautenden [t] ist jedoch im vorliegenden Material unsicher.³

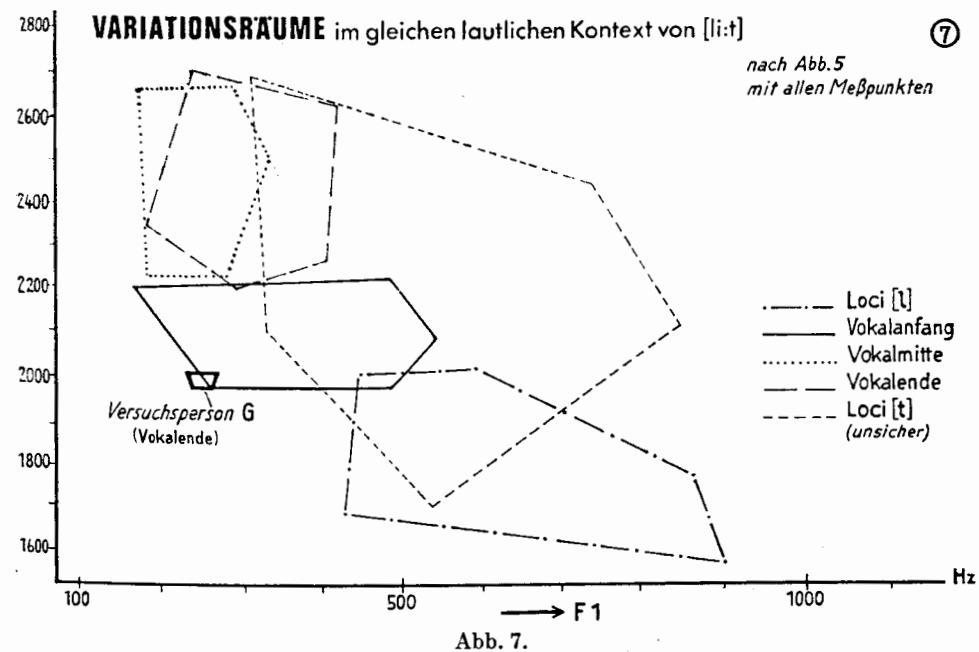


Abb. 7.

³ Das kommt auch in anderen spektrographischen Untersuchungen vor. Vgl. z. B. P. S. Green, Consonant-Vowel Transitions (= Travaux de l'Institut de Phonétique de Lund II). Lund 1959. S. 41.

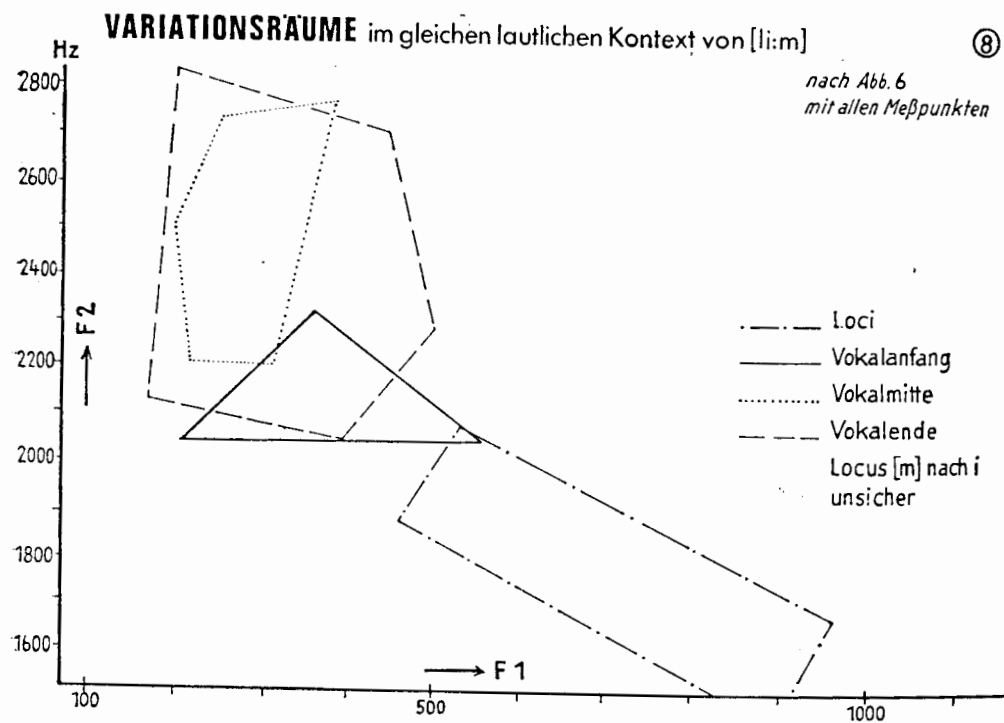


Abb. 8.

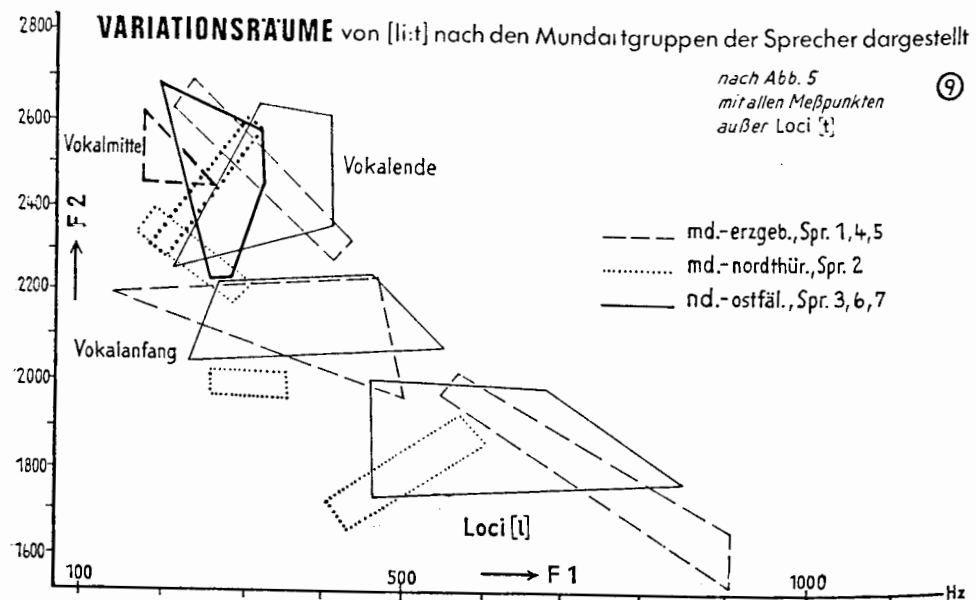


Abb. 9.

Die einzelnen Variationsräume sind bei [li:t] für Locus, Vokalbeginn, sowie Vokalmittle/Vokalende ziemlich scharf abgegrenzt (Abb. 7). Bei [li:m] liegen Vokalbeginn und Vokalende, sowie Vokalmittle und Vokalende ineinander, aber nicht Vokalbeginn und Vokalmittle (Abb. 8). Damit wird die bekannte Tatsache bestätigt, daß im Zeitablauf die Mittle des Vokals die spezifischen Frequenzdaten zeigt und die Messungen dort vorzunehmen sind. — Die Variationsräume, nach den Mundartgruppen der Sprecher für [li:t] dargestellt (Abb. 9), erheben die Forderung nach einer Erweiterung der Einzelbeispiele, die dann vielleicht Hinweise auf Differenzierungsmerkmale erbringen könnten.

Weitere Schlußfolgerungen zu ziehen, verbietet die schmale Materialbasis in Form von [li: + Kons.]. Jedoch sind im Laufe der Untersuchungen Ansatzpunkte für eine Weiterarbeit gefunden worden, die noch nicht weiter verfolgt werden konnten. Das betrifft z. B. die Lage der Loci für F 2, die Dauer der Übergänge vom Locus zum Vokal oder den Frequenzanstieg vom Locus zum Vokal. Ergebnisse versprechen auch die Errechnungen der Distanzen innerhalb der Meßdaten für die Sprecher untereinander. Doch muß das Material überhaupt erweitert und auf andere Lautkombinationen ausgedehnt werden.

КОНСОНАНТНЫЕ ДИФТОНГИ [čĩ], [ǧĩ] В МОЛДАВСКИХ ГОВОРАХ

(Акустические и воспринимаемые характеристики)

Г. М. ГОЖИН*

Сочетания звуков [čĩ] [ǧĩ] в литературном молдавском языке образуют дифтонги консонантного типа, состоящие из двух элементов: палатализованных согласных [c^l], [ǧ^l] и редуцированного [ĩ].

Поскольку в теории молдавского языка отсутствует единое мнение об акустической и артикуляционной природе консонантных дифтонгов, постольку автор доклада предпринял их экспериментально-фонетическое изучение, используя методику и аппаратуру, разработанные в ЛЭФИПР 1 МПНИИЯ. Было произведено их комплексное изучение, включающие методы: рентгенографический, спектрографический, интонаграфический и палатографический, а также метод слухового анализа аудиторами, для которых молдавский язык был родным.

1. ХАРАКТЕРИСТИКА КОНСОНАНТНОГО ДИФТОНГА

а) Литературное произношение

Полусмычной антеропалатальной ротовой звонкой [ǧĩ], по данным комплексного анализа, является консонантным дифтонгом, который образуется в двух артикуляционных фазах. Первая фаза смыкания кратковременна, вторая длительна. В первой фазе звуковая энергия распределена в $F_1—F_3$ и образует диффузный спектр, имеющий общие черты с акустической структурой согласного [g], во второй — в двух высокочастотных формантах, частотная характеристика которых аналогична более диффузному спектру неслогового [ĩ]. При прослушивании отрезков речевого потока, соотносенными с фазой смыкания консонантного дифтонга [ǧĩ] воспринимается четко звук (ǧ), а при аудировании отрезков аффрикатной фазы дифтонга воспринимается неслоговой (ĩ). Дифтонг (ǧĩ), в отличие от полусмычного (ts) характеризуется нечетким переходным процессом.

* ЛЭФ ИЯИЛ Молд. ССР, Кишинев.

б) Северное и южное произношение

Консонантный дифтонг ($\check{g}i$) в северном и южном произношении не отличается от артикуляционной и акустической структуры соответствующего дифтонга в литературном произношении. Спектр образует структуру, состоящую из $F_1—F_5$. Диапазон низкочастотных формант узкий, высокочастотных — широкий. Количество звуковой энергии в высоких формантах уменьшается. Это обусловлено уменьшением объема ротового резонатора.

Переходный период между согласным (\check{g}) и авокальным (i) характеризуется большой длительностью.

в) Центральное произношение

Консонантный дифтонг ($\check{g}i$) в центральных районах МССР переходит в (\check{z}). Согласно рентгенологическим данным фаза смыкания щелевого (\check{z}) становится почти одной продолжительной аффрикативной фазой. Резонансные частоты распределены по всему спектру, амплитуда интенсивности минимальна, частота F_1 смещается в сторону более высоких частот, уровень высокочастотной форманты снижается. Изменение в акустической структуре звука обусловлено сильным уменьшением (сужением) величины ротового резонатора. Таким образом, дифтонг ($\check{g}i$) становится щелевым переднеязычным, палатальноальвеолярным палатализованным согласным (z).

г) Произношение в говорах Каменского и Рыбницкого районов МССР

В этих говорах дифтонг ($\check{g}i$) превращается в ффрикативный щелевой переднеязычный альвеолярно-палатальный мягкий звонкий согласный (z''). При произнесении (z'') отсутствует разграничение между ротовым и гортанно-глоточным резонаторами. Поэтому огибающая спектра приобретает расплывчатую форму с двумя небольшими максимумами, расположенными в область высоких частот. При прослушивании отрезков речевого потока, соотношенными с речевым сигналом, воспринимался согласный (z''). Тональная характеристика (z'') очень повышена.

2. ХАРАКТЕРИСТИКА ДИФТОНГА ($\check{c}i$)

а) Литературное произношение

В литературном молдавском языке сочетание ($\check{c}i$) состоит из антеропалатального полусмычного ротового глухого согласного (\check{c}), образованного в фазе смыкания и авокального (i), который характеризуется аффрикатной артикуляционной фазой. Согласный (\check{c}) характеризуется диффузным спектром, охватывающим $F_1—F_5$, звуковая энергия которого сконцентрирована

в высокочастотных областях. Акустическая характеристика (i) дифтонга ($\check{c}i$) аналогична структуре соответствующего звука консонантного дифтонга ($\check{g}i$).

б) Центральное произношение

В говорах центральной части МССР консонантный дифтонг ($\check{c}i$) становится простым щелевым антеропалатальным альвеолярнопалатальным глухим палатализованным согласным (\check{s}). При произнесении данного звука ротовой резонатор максимально уменьшается, гортанно-глоточный расширяется. Звуковая энергия сосредоточена в области $F_1—F_4$, диапазон низкочастотных формант очень широкий, высокочастотных — узкий. На слух звук (s) характеризуется более низким тоном, чем консонантный дифтонг ($\check{c}i$), который воспринимается как более высокий по тону.

в) Произношение в говорах Каменского и Рыбницкого районов

Консонантный дифтонг ($\check{c}i$) в говорах Рыбницкого и Каменского районов МССР превращается в щелевой переднеязычный альвеолярно-палатальный глухой и мягкий согласный (s''). Он характеризуется малым объемом ротового и большим объемом гортанно-глоточного резонаторов. В спектре звука (s'') обозначаются четыре формантные максимумы, расположенные в областях высоких частот, диапазон которых является очень широким.

Таким образом, данные комплексного анализа свидетельствуют о том, что сочетания ($\check{c}i$), ($\check{g}i$) в говорах молдавского языка образуют консонантные дифтонги. В некоторых говорах эти дифтонги трансформируются.

PHONÉTIQUE ET STRUCTURALISME

MARCEL DE GRÈVE

Je ne suis pas, à proprement parler, phonéticien. C'est donc en intrus que je me présente devant vous. Et, d'emblée, je vous prie de vouloir bien excuser mon outre-cuidance.

Vous aurez compris que l'objet de cette communication sera fort limité, puisqu'il ne pourra évidemment pas être question de faire état de découvertes dans un domaine qui n'est qu'accessoirement le mien. Mes propres recherches linguistiques ressortissent davantage à la logique (qui est pour une bonne part, — faut-il le dire? — l'étude d'un langage particulier) et, bien entendu, aux implications linguistiques de la logique. Toutefois, l'intérêt d'études bénéficiant de l'épistémologie de sciences différentes est, je crois, trop évident et trop bien connu pour qu'il faille en faire encore l'exégèse.¹

Je me propose d'évoquer ici, à la lumière d'un seul exemple, les nombreux problèmes que la linguistique structurale et les analyses logico-linguistiques posent aux recherches phonétiques.

La théorie des „marques distinctives“, ou „traits distinctifs“² élaborée par Troubetzkoy et par Hjelmslev, et exploitée par la plupart des phonologues dans la spécification du phonème, a été, sinon mise en doute, tout de même discutée dans certains de ses aspects considérés comme trop absolus par quelques linguistes.

Ces critiques ou discussions se manifestèrent entre autres à propos des applications psychologiques que Jakobson, Fant et Halle firent de cette théorie, en établissant une corrélation entre, d'une part, l'aspect matériel ou physique des signaux sonores et, d'autre part, les concepts psychiques auxquels ils répondent (point de vue du sujet parlant) ou qui y répondent (point de vue de l'interlocuteur). Je rappelle que, selon eux, le cerveau réagirait de façon „digitale“: en présence de chaque

¹ Léo Apostel remarque très pertinemment que „le théoricien pourra... fournir aux spécialistes les éléments d'une critique interne qui permet de juger progressifs ou régressifs tels ou tels développements“ (*Epistémologie de la linguistique, ds Logique et connaissance scientifique*, publ. p. Jean Piaget, Paris, Gallimard, 1967; p. 1056).

² Angl. "distinctive sound features".

irritation sonore, le cerveau déciderait si le son est labial ou non, explosif ou non, sonore ou sourd, etc.³

D. B. Fry, en se basant sur des résultats de test synthétiques, nia cette „compartimentation“: „(The theory of primary recognition as a result of binary choices) never states quite clearly on which side of the borderline between psychology and physics it operates (...). This (theory) presupposes a one-to-one correspondance⁴ between physical facts and psychological perceptual phenomena which is contradicted by experimental evidence.“⁵ Quant à André Martinet, s'il admet l'utilité de définir les oppositions phonologiques sur la base des différences phonétiques, il met en doute l'application de la qualité de „marque distinctive“ à toutes les différences phonétiques. Il constate, en effet, qu'un certain nombre de variantes sont non pertinentes, ne fût-ce que parce qu'elles sont imposées par la nature, — comme c'est le cas, par exemple, du fait qu'en français les voyelles finales sont normalement brèves, ce qui constitue un phénomène arbitraire et donc non distinctif.⁶

Toutefois, s'il est effectivement fréquent que la théorie phonologique et structurale doive être amendée ou corrigée à la lumière des expériences concrètes de la phonétique, il est non moins fréquent que l'efficacité de l'expérience soit fonction de corrections suggérées par la théorie.

C'est ainsi que la linguistique expérimentale a permis de confirmer la théorie selon laquelle une différence de son ou de timbre n'entraîne pas nécessairement une différence de signification, ni pour le sujet parlant, ni pour l'interlocuteur. A cette constatation s'ajoute la découverte, au moyen d'appareils acoustiques, qu'il n'y a pas deux sons absolument identiques: un phonème x , par exemple $[a]$, „émis“ par m sujets parlants à peu près de i façons différentes, c'est-à-dire avec i fréquences différentes entraînant des données phonétiques différentes, sera „enregistré“ par n oreilles humaines de j façons différentes, entraînant l'enregistrement de ij fréquences, soit:

$$(a) \quad (\forall x) [f(x_i, y_j)].$$

On constate, en outre: 1° que le même sujet parlant m^1 et le même auditeur n^1 émet ou enregistre des fréquences différentes selon le moment; 2° que le phonème émis x ni le phonème reçu y n'ont de spécificité immuable, mais qu'ils se définissent par

³ Voy. Roman Jakobson, Gunnar Fant et Morris Halle, *Preliminaries to Speech Analysis. The Distinctive Features and Their Correlates* (Cambridge, Mass., University Press, 1952) et Roman Jakobson et Morris Halle, *Fundamentals of Language* (La Haye, Mouton, 1956).

⁴ C'est-à-dire une relation bi-univoque.

⁵ D. B. Fry, *Speech and Language*, ds *The Journal of Laryngology and otology* (t. LXXXI, 1957, pp. 434—452), p. 442.

⁶ André Martinet, *Economie des changements phonétiques. Traité de phonologie diachronique* (Berne, Francke Verlag, 1955), p. 18. Voy. aussi H. Frei, *Langue, parole et différenciation*, ds *Journal de psychologie* (1952), pp. 140—149.

rapport à leur entourage (phonétique, physique, psychologique, etc.) sous forme de relation fonctionnelle. Il s'ensuit qu'on assiste à un ensemble disjonctif continu de conjonctions continues, soit:

$$(b) \quad \sum_{ij} \{ \prod_j [f(x_{ij}, y_{ij})] \}^7$$

Or, nous savons, — pour n'envisager qu'un seul cas, — que dans une suite conjonctive continue la fausseté d'un seul élément entraîne la fausseté de l'ensemble. Et pourtant, nous ne savons que trop bien que si le sujet parlant émet et si l'auditeur enregistre autant de „sons“ différents, il n'en demeure pas moins que chacun „dit“ ou „entend“ autant de fois „la même chose“. Ce qui signifie que, contrairement aux résultats des recherches expérimentales, la formule précédente (b) $\equiv 1$.

Il me semble que ce fait doit mettre le phonéticien de stricte observance dans l'embarras et placer la phonétique proprement dite devant un mur infranchissable, — confirmant ainsi que la phonétique ne peut s'en tenir exclusivement à la description et à l'analyse de la morphologie externe, sans tenir compte de la morphologie interne qui est faite de rapports d'isomorphisme et de relations fonctionnelles ou relations entre éléments de classes (ou ensembles) différentes.

Par contre, ce même fait représente pour le phonologue, non seulement une preuve (supplémentaire) de la combinaison fonctionnelle des phénomènes linguistiques, mais il constitue en outre la base éventuelle de nouveaux progrès analytiques: le phonologue pourra découvrir des éléments qui doivent lui permettre des recherches plus fouillées.

En effet, la formalisation ayant abouti à l'expression (b) précitée permet d'introduire la théorie des „traits distinctifs“, „arbitraires“ inclus, dans un calcul logique, — ce qui est une condition suffisante et nécessaire pour 1° vérifier l'exactitude (ou la correction) d'un raisonnement à partir de ces données, 2° établir des recherches statistiques avec ces éléments comme base, ou encore 3° leur appliquer un calcul de probabilité. Serait-ce, en outre, vraiment présomptueux de considérer qu'il n'est pas exclu qu'un calcul logique conduit de cette manière puisse éventuellement montrer au chercheur la voie à suivre?

Quoi qu'il en soit, si les appréhensions d'André Martinet ont été effectivement confirmées, elles se présentent dans le même temps sous un tout autre jour: les variantes qu'il a présentées comme étant „arbitraires“ le deviennent beaucoup moins, déterminées qu'elles sont par des relations fonctionnelles authentiques, mais non encore entièrement définies.

⁷ Pour rappel:

$$x_1 \wedge x_2 \wedge x_3 \wedge x_4 \wedge \dots \wedge x_i \rightarrow \prod_i x_i$$

$$x_1 \vee x_2 \vee x_3 \vee x_4 \vee \dots \vee x_i \rightarrow \sum_i x_i$$

TRANSIENT PHENOMENA IN MUSIC AND SPEECH

L. M. GROBBEN*

INTRODUCTION

Transients are known to contribute to the recognition of a typical sound. Sound signals become unrecognizable if their beginning and end are smoothly suppressed. Many investigators have examined the transient phenomena; e.g. Stumpf (1910), Backhaus (1932) Trendelenburg and Franz (1935) and others all came to the aforementioned conclusion with the aid of octave filters, oscillographs and a planimeter. They measured onset times and growth of the harmonics as important properties for speech and musical instruments. They stated a.o. that the onset time of flute and violin was rather long (100—200 m s) and of e.g. trumpet rather short (15 m s); the clarinet would have an intermediate position.

Recent investigations of a more psycho-acoustic character (Saldanha and Corso [1964] and Berger [1963]), deal with the importance of attack, steady state, release and vibrato of musical sounds in a rather restricted area of the musical frequency spectrum.

Our investigations deal with acoustical analysis with modern instrumentation and with psycho-acoustic evaluation as well of the aforementioned transient and steady state phenomena of musical instruments, with the intention for future research on speech. We repeated the older investigations with regard to the growth of amplitudes of harmonics in a series of specific sounds, produced by violin, double bass, flute, oboe, clarinet, bass-clarinet, bassoon, horn, trumpet, trombone.

METHOD

The sounds to be analysed were taken from phonographic record and equipment (Lenco L 77) or by live recording with an electrodynamic microphone (Sennheiser HN 421) recorded on tape (Revox G 36; Scotch tape 203). Often an infinite tape loop was used. Exploration was performed with an 1/3 octave filter (Bruel & Kjaer 2112)

* Lab. f. Labyrinthology (Head Prof. J. J. Groen) of the O. R. L. Dept State University Utrecht (Dir. Prof. Dr. P. G. Gerlings).

and a narrow band filter (Bruel & Kjaer 2105); the onset of the sound signal was reproduced on the screen of an oscilloscope (Tektronix 561 A) and photographed. Remark: The build-up time of the filter appeared to be negligible.

On the second track of the tape loop, a pilot pulse was recorded to mark the beginning of the sound in order to synchronize all the analyses pertaining to the same signal. A 100-cycle tone was reproduced as a time marker on the oscilloscope screen via the second beam. A Panoramic Sonic Analyzer (model A P-1) and a Kay Sona-Graph (6061 A) were used to obtain an overall spectrogram and a control of the relative importance of the harmonics.

Fig. 1 represents the blockdiagram of the apparatus.

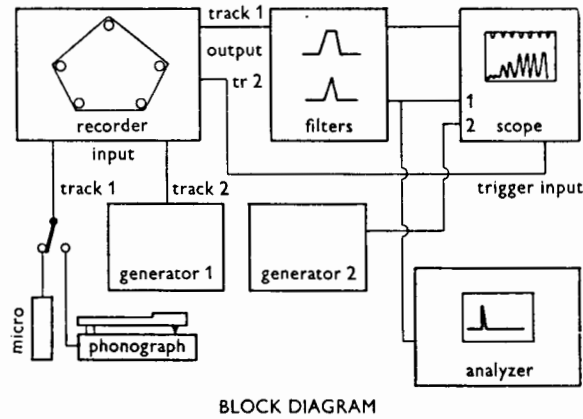


Fig. 1. Blockdiagram of the apparatus.

Fig. 2 gives an example of an analysis of the onset as a function of time (flute, pitch 880 cps). The arrow marks the moment where the onset is considered to be ended.

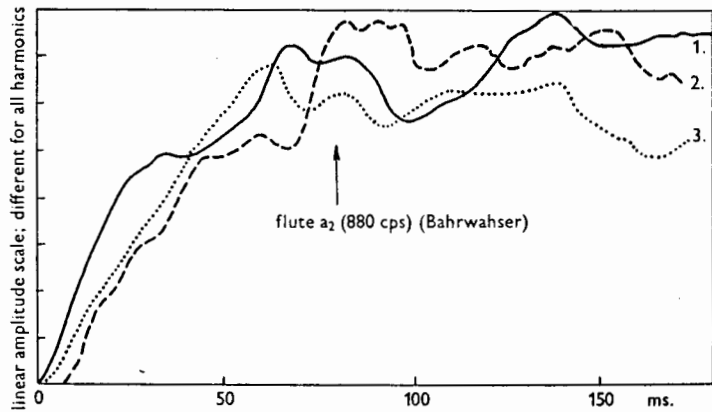


Fig. 2. Example of an analysis of the onset as a function of time. (Flute, pitch 880 cps). The arrow marks the moment, where the onset is considered to be ended.

RESULTS

There appears to be a specific relation between the growth of the harmonics and the pitch of the fundamental of a specific instrument. In first approximation, the growth is inversely proportional to the frequency of the fundamental; in other words:

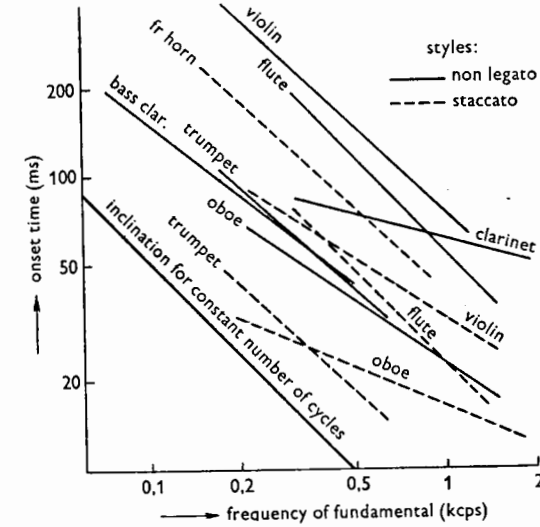


Fig. 3. Onset time as a function of frequency for different musical instruments, with the style as parameter.

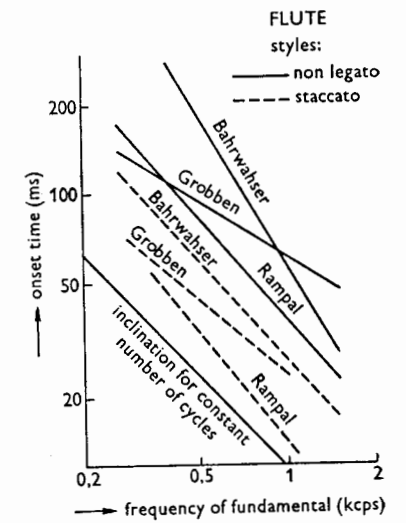


Fig. 4. Onset time for the flute for different players.

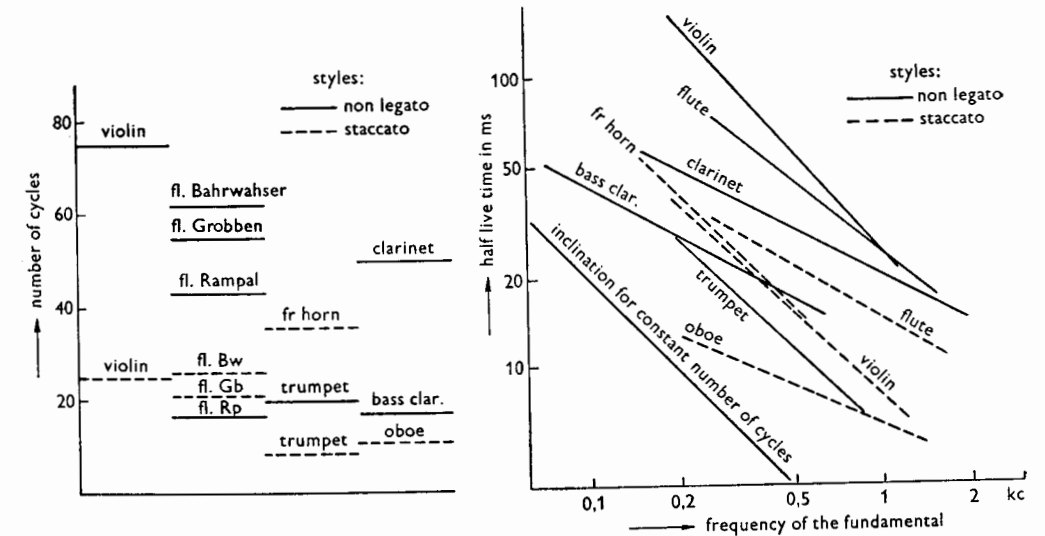


Fig. 5. Number of cycles necessary for the completion of onset for different instruments and styles.

Fig. 6. Half-life time of the earliest harmonic.

the number of oscillations necessary for building up to steady state is constant for a specific instrument, with an additional influence of the artist's ability and style. There are, however, some deviations.

Fig. 3 gives the measurements for different instruments as a function of fundamental, with the style as parameter.

Fig. 4 stresses the importance of different players and style for one instrument (flute). The flutist Bahrwahser of the German School builds up his tone markedly slower than the French player Rampal; the author, amateur-flutist occupies an intermediate position as becomes a Dutchman.

Fig. 5 same data as in fig. 3 but now arranged according to the number of cycles necessary for a completion of onset for different instruments and players.

Fig. 6, 7 and 8 represent these same measurements again, but now from the point of view of half-life time of onset or appearance of the earliest harmonic.

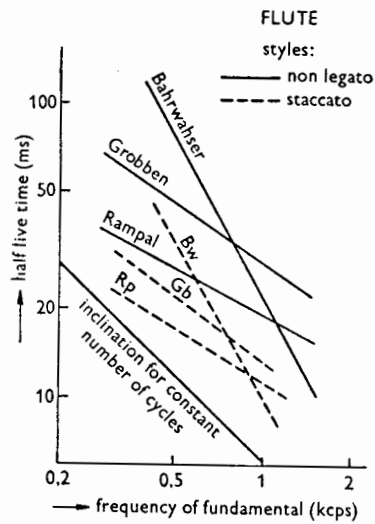


Fig. 7. Half-life time of the earliest harmonic for different flute players.

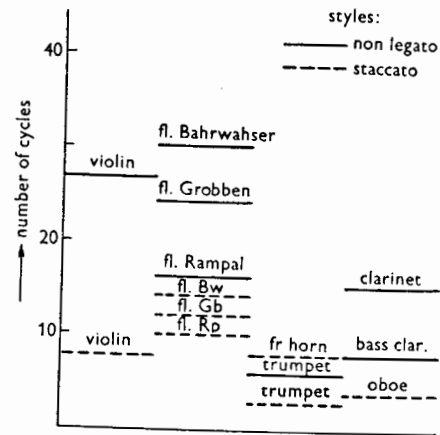


Fig. 8. Number of cycles necessary for reaching half-life time of the earliest harmonic. Parameter: instrument and style of playing.

DISCUSSION

Although Backhaus mentions the relation between pitch and onset-time, especially in the lower frequency range, up till now the fixed relationship between necessary number of cycles for the production of the tone and its pitch (depending further upon instrument, player and intention) has never been stated, as far as the present author knows.

It has to be mentioned here, that the present author has used as a criterion for decision whether the onset of the sound is completed, the behaviour of the strongest

partials, neglecting all partials, which are 20 dB or more weaker than the strongest one. According to our opinion, this seems to be a reasonable procedure, because: a) our results are reproducible within reasonable limits, b) they are roughly in agreement with the results of Backhaus and others, c) recognition of instrument seems to be determined for the major part by the appearance in the onset of the strongest partials, as has been verified by tentative investigations where an experienced team of listeners gave their opinion. The sound to be judged was given in a series of presentations of successive filtering; starting with presenting merely the fundamental. All measurements will be continued and extended e.g. in the analysis of the importance of envelope form. In a later stage speech will be analysed according to the procedure described.

REFERENCES

1. Stumpf C. *Die Sprachlaute*. 1926 Springer Berlin p. 374.
2. Backhaus H. Über die Bedeutung von Ausgleichsvorgängen in der Akustik. 1932, *Z. s. f. techn. Physik* 13 p. 31.
3. Trendelenburg F. u. Franz E. Untersuchungen an schnellveränderlichen Schallvorgängen. 1935, *Z. s. f. techn. Physik* 16 p. 513.
4. Saldanha E. L. and Corso J. F. Timbre cues and the Identification of Musical Instruments. 1964, *J. A. S. A.* 36 p. 2021.
5. Berger K. Some Factors in the Recognition of Timbre. 1964, *J. A. S. A.* 36 p. 1888.

DISCUSSION

Filip:

I would like to ask whether also instruments of different quality but of the same type were investigated. If so, what results were obtained.

Fransson:

The formants of the double-reed wood-wind instruments are of great importance for the characteristic timbre of these instruments during steady state.

The formants of the speech vowels belong to the vocal tract but these formants are produced by the reeds, i.e. they belong to the source and not to the instrument. I think it would be of interest to know the building-up process of the formants during the transient period and investigate the influence of this process on the perception.

Sovijärvi:

The human articulatory organ as a musical instrument (using singing voice) has doubtless a very complex onset system. I would like to mention only one case: the vowel [a] after a voiceless plosive, for example in [pa, ta, ka], has after noise of the explosion an onset time of 10–20 msec. measured by means of contour spectrograms of the Voiceprint Spectrograph.

Grobbe:

Ad Filip: I measured this only for the flute; Figure 4 and 6, but not with the particular intention of making a comparison. Lottermoser and others made these measurements for violins. I intend to do the same for other musical instruments.

Ad Fransson: I thank you for your comment. I certainly will think it over and carry out some measurements. Perhaps it is really better to measure formants than harmonics for particular musical instruments.

Ad Sovijärvi: I performed preliminary measurements on speech too and I found longer onset times. I am not yet sure about the measurements however.

INTONATION IN FRENCH

R. GSELL

DISCUSSION

Rossi:

J'ai été très intéressé par votre communication. Je me permets cependant la remarque suivante. Vous dites que l'intonation a une fonction d'actualisation; vous attribuez la fonction démarcative à l'accent; pour ce faire vous devez identifier linguistiquement l'accent. Par quelle procédure réussissez-vous à le dissocier l'intonème de groupe?

Spang-Thomsen:

Dans le petit article auquel M. Gsell a bien voulu renvoyer,¹ j'ai partout donné deux mesures alternatives pour la quantité des voyelles finales accentuées. En effet, je n'osais pas affirmer si les vibrations très faibles sur lesquelles finissent les tracés jouent le rôle pour l'impression auditive, si elles font, par conséquent, partie de la voyelle. Si oui, la durée longue semble être partout une manifestation de l'accent. Si non, l'accent ne semble pas nécessairement se manifester par une prolongation de la voyelle. Je comprends que M. Gsell a opté pour la première solution et je voudrais demander quelles sont ses raisons de ce choix.

Горсуева:

Я согласна с Гселлем, что интонация это дискретный феномен. По моему, она имеет два типа единиц: интономы и супраинтономы, состоящие из интоном. Интономы — не прямые знаки, супраинтономы — прямые. Интономы состоят из компонентов интонации, взятых в максимальном значении в рамках предложения. Интонация имеет два уровня и представляет собой третье членение языка.

¹ L'accent en français moderne. Quelques expériences sur sa manifestation phonétique — Études romanes dédiées à Andreas Blinkenberg. Copenhague (Gyldendal) 1964, pp. 181—208.

L'IMPULSION SYLLABIQUE ET SA PERCEPTION D'APRÈS LES ANCIENS PHONÉTICIENS ARABES

A. HADJ-SALAH*

La syllabe telle que l'ont définie les Grecs a été connue des Arabes par la traduction des ouvrages de l'Antiquité. Cependant la tradition linguistique antérieure à cet évènement comporte un système très original basé sur les notions de „*ḥarf*“ et de „*ḥaraka*“. Le *ḥarf* est, d'après les grammairiens arabes (cf. Rummani, *Commentaire du Kitab*, Manuscrit microfilmé, Ligue Arabe; t. 5, f. 41v) „le plus petit segment de la chaîne parlée“, mais „il ne peut être réalisé isolément car la liaison séquentielle (des segments) est la règle . . .“. Les auteurs d'inspiration hellénistique, eux-mêmes, ont retenu cette notion de *ḥarf* lorsqu'ils essayèrent de définir la syllabe des Grecs: „un segment non sonant qui est dans un état tel qu'on puisse le prononcer dans un continuum naturel est appelé syllabe (*maqṭa'*)“ (Avicenne, *al-Šifā*, Caire, t. III, 123). *Al-Khalīl* (grammairien très célèbre du 8^e s.) attribue au *ḥarf* deux qualités: un *ḡars* qui signifie timbre et un *ṣarf* qu'il explique par *ḥaraka*. Or *ṣarf* a un sens lexical de base qui correspond à la notion physique de force de cohésion. *Al-Khalīl* montre, en effet, que l'élément quantitatif correspondant à la partie finale d'une syllabe longue (et qui est considérée avec raison par les Arabes comme segmental et non parasegmental) ne peut supporter le *ṣarf* de la *ḥaraka* c.-à-d. la force de cohésion que lui applique la *ḥaraka*. Celle-ci serait donc une qualité du *ḥarf* qui lui permet de se produire, autrement dit l'impulsion nécessaire au franchissement de l'obstacle occasionné par le contact organique (*ietimād*) à un point quelconque du canal phonatoire.

Al-Khalīl était aussi un théoricien de la musique. Or nous savons par les traités de musicologie arabes que la rythmique avait pour fondement la notion de *ḥaraka*. *Al-Fārābī* nous dit, en effet: „Les Arabes qualifient de percussion sans motion (= *sukūn*) celle qui est suivie d'une pause, et de percussion avec motion (*ḥaraka*) celle qui n'est pas suivie de pause, mais d'un mouvement vers la note suivante“ (trad. du B. d'Erlanger dans la *Musique arabe*, t. II, 31). Cette conception du rythme correspond exactement aux critères de successivité des sons du langage établis par les Arabes.

Ainsi une *ḥaraka* constitue également le mouvement acousticophysiologique qui doit se faire d'un *ḥarf* vers un autre *ḥarf* pour rendre possible l'acte de parole.

* Institut de Linguistique et de Phonétique, Alger.

Sibawayhi, célèbre disciple d'*Al-Khalil*, nous dit aussi: „On ne peut faire suivre l'articulation d'un *ḥarf* non doté de *ḥaraka* par un son vocal, sinon ce *ḥarf* en serait muni“ (*Kitāb*, Būlāg, II, 285) et „Si le locuteur veut produire les *ḥuruf-s*, il doit élever ou baisser la voix, selon son désir, en réalisant les *ḥurūf-s* al-madd (segments quantitatifs) ou seulement les éléments qui en dérivent (les *ḥarakāt-s*)“ (II, 406), enfin „le locuteur ne peut doter d'une *ḥaraka* ce après quoi il veut garder le silence“ (II, 279).

La *ḥaraka* est donc pour cet auteur le mouvement qui actualise le *ḥarf* en le plaçant dans un enchaînement sonore.

Il est intéressant de remarquer que les auteurs arabes n'ont pas donné de nom à la syllabe brève mais qu'ils possédaient une dénomination pour la syllabe longue. Pour les anciens théoriciens arabes, comme pour nous, il est impossible de prononcer une syllabe brève (ouverte) à l'état isolé ou à la pause: une telle syllabe est nécessairement longue (plus exact. fermée). en ce sens que la voyelle qui la termine est nettement décroissante et extensible (cf. M. Durand, *Voy. longues*, 177). Si c'est une consonne qui ferme la syllabe, elle est évidemment implosive. La séquence C + voyelle brève n'a pas d'autonomie: la voyelle brève implique le passage à un autre phonème. Le sentiment qu'a le sujet parlant de l'existence d'unités successives vient de ce qu'il transforme ces combinaisons, par la lenteur du débit, en syllabes longues non susceptibles d'être liées. Dans un débit normal la seule unité réelle mais *imperceptible* est l'impulsion aéroorganique provoquée par les contractions rapides des muscles intercostaux (G. Stetson, *Motor phonetics*). *La succession de ces impulsions motrices est au continuum physiologique ce qu'est la succession des impulsions sonores de la glotte au continuum sonore*. Il n'y a pas entre elles des limites telles qu'on puisse les considérer comme des unités autonomes.

On serait tenté de voir alors dans la *ḥaraka* du *ḥarf* (en tant que qualité) le caractère explosif des consonnes ou même croissant des segments de la parole. En fait la *ḥaraka* a un contenu beaucoup plus simple et aussi plus large. Je proposerais (avec al-Khalil, s'il avait été notre contemporain) le terme de *kinème* pour l'impulsion syllabique, de *kinème zéro* pour l'absence d'impulsion, de *kinèse* pour l'état de *ḥaraka* et d'*akinèse* pour son contraire. Les segments peuvent alors être *kineisés* ou *akineisés*.

Pour délimiter les contenus de ces termes, il est nécessaire de distinguer trois niveaux:

I. *Un niveau générateur ou physiologique* qui constitue le substrat matériel de la parole. Il présente deux aspects:

— *l'aérokinèse* (*ḥarakat al-hawa'*), élément véhiculaire, le plus petit mouvement de l'air phonateur étant l'aérokinème.

— *l'organokinèse* (*ḥarakat al-'uḍw*), élément moteur; le plus petit élément organique donnant naissance à un son linguistique.

II. *Un niveau transmetteur ou acoustique* représenté par la *phonokinèse* (*ḥarakat al-sawt*), le plus petit élément sonore ou chuché pouvant accompagner un kinème.

III. *Un niveau linguistique* proprement dit. C'est la résultante de toutes les

composantes kinétiques. Les mouvements aéroorganiques ont des zones et des modes de production distincts d'où les différents timbres vocaliques. En arabe, trois timbres kinémiques à fonction phonologique ont été relevés: damma (u) fatha (a), kasra (i).

La perception du kinème, pour les Arabes, ne se fait qu'au niveau acoustique. Les soubassements aéroorganiques passent inaperçus. On s'est demandé au 4^e siècle de l'Hégire si le phonokinème était perçu avant, après ou en même temps que le segment kineisé. C'était confondre les niveaux sus-indiqués. La faute en est due à des grammairiens qui n'avaient pas très bien saisi cette différence et aussi à l'influence grecque à laquelle on avait emprunté les notions d'essence et d'accident. Certains auteurs avaient ainsi considéré le *ḥarf* comme une essence et la *ḥaraka* comme son accident. Mais de nombreux auteurs ont montré que la force de retention (*ḥabs*) qui aboutit au kinème se situe avant le segment kineisé, que le mouvement kinémique a son point de départ immédiatement après la rétention c.-à-d. en même temps que l'explosion, qu'enfin le son ou phonokinème est subséquent à la production du segment.

Ce que l'on peut retenir dans ce système c'est sa capacité d'extension à d'autres langues que l'arabe. Les linguistes arabes se sont refusés à voir dans le 1^{er} élément des groupes explosifs un segment akineisé. Ce serait une impossibilité matérielle, disent-ils et non une affaire de convention linguistique. On leur opposait des mots persans tels que „*stām*“ ou „*klūd*“; ils répondaient que le S et le K étaient bien kineisés mais que leur kinème était si bref qu'on ne pouvait apprécier nettement sa qualité acoustique. La différence entre la notion de voyelle et celle de kinème apparaît très clairement: entre s et t, par ex., il y a un kinème mais pas de voyelle.

DISCUSSION

M. Cohen:

Les oeuvres des grammairiens et phonéticiens arabes sont importantes, en particulier par la connaissance de divers mécanismes de la parole. Elles ne dépendent que peu des études des Grecs; d'autre part on peut se demander s'il n'y a pas de lointaines traditions mésopotamiennes qui auraient été continuées dans l'Inde, d'où certaines notions ont pu passer aux Arabes. Il est très heureux de voir l'oeuvre des grands savants arabes reprise et continuée par des savants qui sont d'autre part au courant de tout l'acquit de la science européenne.

PSEUDO-UNITS IN PHONETICS

GÖRAN HAMMARSTRÖM

O. Descriptions of spoken languages have mostly been based on materials which a linguist, knowing the language, has established by "rapid" listening "directly" to speakers or "indirectly" to recordings of speakers. Details that had not been actually heard were added by the linguist. Often the linguist could use his own knowledge of a language without even listening to any speaker at all. This technique seems to be practical and allows for establishing the units usually (and rightly) considered to be basic (" α -units" such as phonemes, morphemes etc).

Details belonging to what I have called the β -level, expressing how an utterance is said, and the γ -level, characterising the speaker, (cf. my *Linguistische Einheiten*..., Berlin—Heidelberg—New York 1966, p. 8—13) generally require painstaking "indirect" listening to tape recordings where many repetitions are often necessary to establish each detail, i.e. "rapid" listening is not sufficient in this case. In addition, the listening tests that would be required for many problems have not yet been undertaken. Consequently our knowledge of these levels is in some respects incomplete.

1.0. Little interest has been given to what I would like to call pseudo-units. Sometimes parts of the "spoken chain" (the expression) are such that they seem to be "realisations" of well known α -units such as phonemes, syllables or prosodemes but in the usual description they cannot be related to any such units. The explanation of this fact is that the "rapid" listening of the normal linguist is directed exclusively towards clear, rather slow and unemotional pronunciation. This procedure provides an acceptable starting point but is not the basis for the most complete possible description.

I would now like to consider a number of pseudo-units.

1.1. *Pseudo-phones*. The segments of the spoken chain are not always in one-to-one relation to phonemes. The Portuguese word *geral* is sometimes, under emphasis, pronounced in a way that could be written *gerali*, i.e. at the end of the word there appears an "extra" segment of the type that from the viewpoint of the sound can usually be attributed to a phoneme although in the word mentioned this does not seem meaningful. When the word written *geral* is described, it is said to end in an $[l]$. We suggest that a segment such as the final $[i]$ of "gerali" be called a pseudo-phone.

Even if a segment such as the one just mentioned is not what it might superficially seem to be, i.e. a segment belonging to a phoneme, it has a function on the α -level being part of a possible realisation of a "longer" unit than a phoneme, namely a lexeme. Thus it can be treated less on the phonological than on the morphological level. At the same time it is certainly part of a β -unit. (In the example given the β -value is emphasis.) The Swedish so called supradentals, such as [s] in *mors* (mother's) are from an auditory, acoustic and articulatory standpoint comparable to phones. In the phonemic description that I prefer, however, each supradental does not correspond to a phoneme and one might be tempted to call them pseudo-phones. As, on the other hand, each supradental corresponds to *two* consecutive phonemes ([mo:s] = = /mo:rs/), it is preferable not to do so. A similar, but not altogether equal, case is constituted by a segment that occasionally (and not normally like a Swedish supradental) realises two phonemes. "In rapid 'uncareful' speech, for example, an English vowel with a following [n] may occasionally fuse to a single-segment portmanteau nasalized vowel [...]" (Pike, K. L., *Language in relation to a unified theory of the structure of human behavior*, II, 1955, p. 16). This type of segment, too, would not be considered as a pseudo-phone (but as an "unclear segment", see below, paragraph 4).

1.2. *Pseudo-syllabs and pseudo-syllab prosodes*. The Portuguese word *geral* is generally considered to possess two "syllables" (syllabemes) but in the emphatic pronunciation just mentioned it has three "syllables": *ge-ra-li*. Of these three "syllables" the first is clearly a syllab (because it can be attributed to a syllabeme). The third is clearly a pseudo-syllab not being attributable to a syllabeme. The second, too, must be considered as a pseudo-syllab being pronounced in a way that corresponds to the letters *ra* whereas the syllab corresponds to *ral*.

Pseudo-syllabs can be described in the part of the description dedicated to syllabemes and syllabs but they are above all relevant on the α -level as parts of lexes. They are at the same time parts of β -units (expressing e.g. emphasis).

Syllabs are characterised, i.e., by syllab prosodes which are attributable to syllabeme prosodemes. Pseudo-syllabs are characterised by pseudo-syllab prosodes that cannot be attributed to syllabeme prosodemes. (Cf. *Linguistische Einheiten...*, p. 37—40.)

1.3. *Pseudo-syntagms and pseudo-syntagm prosodes*. The usual syntactic analysis is based on clear, unemotional speech (which is one of the range of possibilities of the β -level). The resulting syntagms are pronounced and heard as groups of lexes. Under determined β -conditions the lexes are grouped in a way that does not correspond with the usual syntagms of the sentence. These groups of lexes can be called pseudo-syntagms. Unlike syntagms they cannot be attributed to syntagmemes (cf. *Linguistische Einheiten...*, p. 46).

Pseudo-syntagms can split up the syntagms pronounced and heard in most cases or they can join together lexes usually belonging to separate syntagms, i.e. determined syntactic means, that usually group together the lexes of a determined syntagm,

are used by the speaker to group this syntagm and one or more adjacent syntagms into a "long" pseudo-syntagm. An example of the second possibility is offered by a recording that a French speaker made for me some years ago. The three consecutive sentences *Il est trop petit. Ce n'est pas pour lui. C'est pour les grandes personnes* are usually pronounced so that the end of each sentence is clearly indicated (i.e. through falling pitch) but by the speaker mentioned the three sentences were grouped into one pseudo-sentence. This was indicated i.e. by rising pitch at the end of the first two sentences. The result is a stylistic (β -) effect: The three sentences become closely tied together. The first two sentences enumerate two closely related facts and the third closely related fact of the third sentence is added in a particularly conclusive way.

Syntagms are characterised by syntagm prosodes belonging to syntagmeme prosodemes but the pseudo-prosodes of pseudo-syntagms do not correspond to such prosodemes.

It should be noted that, although syntactic analysis does not usually concern stretches longer than a sentence, in the French example just mentioned three sentences can be said to form one pseudo-sentence.

2. It seems doubtful to what degree the notions of pseudo-lex and pseudo-lex prosode would be interesting and useful.

3. Common to the stretches of sound corresponding to the pseudo-units we have been concerned with is:

a) They are segments of speech which usual linguistic description, being based on clear, "neutral" speech (or on the thoughts the linguist has about such speech), does not take into account. In a description, where the facultative variants of α -units are included, they will, however, necessarily be considered.

b) They have different β -functions, being the expression of β -units or parts of the expression of such units. (β -units can usually be described as facultative variants within α -units.)

4. It should be added that in a complete description of speech, account must be taken not only of such segments as mentioned above, but also of segments that can be attributed neither to "traditional" units nor to pseudo-units. Examples of such "unclear segments" are, e.g., vocal murmurs of the sort that cannot be attributed to a phoneme. The lex [naəs] as a possible pronunciation for the Swedish lexeme *naturligtvis* would contain at least one such unclear segment, i.e. [ə] (cf. *Linguistische Einheiten...*, p. 44). Thus the description of the segments of a lex must take into account, among other things, (a) phones, (b) pseudo-phones and (c) nuclear segments.

SYSTEMS OF LATERAL SOUNDS AND PERCEPTION

ERIC P. HAMP*

In the brief time we have at our disposal we wish to consider the nature of the distinctions found in languages rich in lateral sounds. Phonologies involving but one lateral and one retroflex can often be accommodated without any mention of laterality in the matrix of underlying distinctive features; e.g. by [\pm continuant], or the like. It has recently been suggested by workers in generative phonology that a feature of *laterality* be incorporated in the list of distinctive features from which human languages choose a repertory. Such a feature is clearly called for in certain languages; the question then arises whether further complexities are found in laterals or whether other well established features will account for the rest of the variety that may be found in a rich set of laterals.

Elaborate sets of laterals are attested in the Caucasus. These were studied by Troubetzkoy (BSL 23.184—204, 1922) at a time when not so much detail was available as is the case today; moreover, much of his argument was concerned to show the genetic correspondences, and to trace the proto-set of relevant obstruents, some of which seem to be surely non-lateral in origin (I would interpret the proto-set, on Troubetzkoy's evidence, as: * $\text{ʃ}, \text{l}, \text{x}, \text{x}'$). Be that as it may, the richest Caucasus sets show at least 3 laterals, and none exceeds 5 (Avar-Andi and Arč'i). Though Ubykh has 3, the glottalized affricate $\text{ʃ}'$ occurs only in $p'\text{ʃ}'\text{ə}$ '4' and in Abzakh loans (Vogt, *Dictionnaire*, Oslo 1963, 18).

Basing ourselves on T. E. Gudava, *Konsonantizm andijskih jazykov* (Tbilisi 1964), we see that Andi has the maximal C system, and Akhvakh the maximum laterals; the latter presumably original for Andi:

* University of Chicago.

	<i>b</i>	<i>d</i>	<i>g</i>	<i>ẓ</i>	(<i>z</i>)	<i>ɣ</i>	<i>ω</i>		<i>b</i>	<i>d</i>	<i>g</i>	<i>ẓ</i>	<i>z</i>	<i>ɣ</i>	(<i>ω</i>)	
lax						<i>q̣</i>	<i>ḷ</i>					<i>k</i>	<i>č</i>	<i>c</i>	<i>q̣</i>	<i>ḷ</i>
asp	<i>p</i>	<i>t</i>	<i>k</i>	<i>č</i>	<i>c</i>				<i>p</i>	<i>t</i>						
tens						<i>q̣^o</i>	<i>ḷ^o</i>					<i>k^o</i>	<i>č^o</i>	<i>c^o</i>	<i>q̣^o</i>	<i>ḷ^o</i>
lax			<i>ḳ</i>			<i>q̣^o</i>	<i>ḷ^o</i>					<i>ḳ</i>	<i>č̣</i>	<i>c̣</i>	<i>q̣^o</i>	<i>ḷ^o</i>
chk		<i>ṭ</i>		<i>č̣</i>	<i>c̣</i>		<i>ʔ</i>		<i>ṭ</i>							(<i>ʔ</i>)
tens			<i>ḳ^o</i>			<i>q̣^o</i>	<i>ḷ^o</i>					<i>ḳ^o</i>	<i>č̣^o</i>	<i>c̣^o</i>	<i>q̣^o</i>	<i>ḷ^o</i>
lax						<i>x</i>	<i>ḷ</i>					<i>š</i>	<i>s</i>	<i>x</i>	<i>ḷ</i>	
spir			<i>x̣</i>	<i>ṣ̌</i>	<i>ṣ</i>		(<i>h^o</i>)	<i>ḥ</i>	<i>x̣</i>					(<i>h^o</i>)	<i>ḥ</i>	
tens						<i>x̣^o</i>	<i>ḷ^o</i>					<i>ṣ̌^o</i>	<i>ṣ^o</i>	<i>x̣^o</i>	<i>ḷ^o</i>	
	<i>m</i>	<i>n</i>			<i>r</i>	<i>l</i>			<i>m</i>	<i>n</i>				<i>r</i>	<i>l</i>	
	<i>w</i>	<i>j</i>							<i>w</i>	<i>j</i>						

Allowing for surface redundancies, I would posit for these the features [voice], [contin], [checked], [tense], [nas], [sonant]; the points of articulation seem to be accommodated most naturally as:

[—compact]	—labial	[+flat]]	(Note the corresponding poor rôle of flatness in Caucasus vowel systems.)
	—dental	[—flat]	
[+compact]	—velar	[—grave]	[—strident]
	—postvel.	[—lat]	
	—lateral	[+lat]	[+grave]
	—silibant	[—grave]	[+strident]
—shibilant	[+grave]		

Grouping by compactness associates those articulations rich/poor in continuance and tenseness; the pair *b/d* matches *m/n* and *w/j*. Grouping by laterality associates the genetic relations observed by Troubetzkoy (see above)—a welcome additional result.

Thus the laterals are specified as [+comp —strid +grv +lat ±contin ±checked ±tense], and [±obstruent] for the sonant. The laterals are thus here maximally specified consonants, but in no way systematically more divergent. Phonetically there are no unusual or unexpected articulations: Apart from tenseness (“gemina-tes”), we have [λ λ' l l].

On the Northwest Coast of North America, in such languages as Quileute and the Salish family, we similarly find proliferated laterals, and the need for a feature of laterality. But beyond that, laterals are simply consonants like any other articulation, and fit within the cadre of glottalization and spirancy that applies generally to these systems.

On comparative grounds it is easy to show that Early Albanian had three laterals, and as a matter of fact I have actually found them in the Arvanitika dialect of Mándra, in Greece: [l l' λ]. Phonetically, the set is unusual, velarized, palatalized, and palatal; the distinction between the last two is very hard to hear, and it is easy to see why all other Albanian dialects have changed the articulations or undergone mergers. But when we recall that Albanian has two retroflexes [r r̄], and nasals including [n ñ], we readily associate *l'* with *r̄*, and *λ* with *ñ*, by assigning them shared features, features that are in any case required.

Old Irish must have had 4 laterals (as it also had 4 nasals, apart from the labial nasals, and 4 retroflexes), but these were easily specified by the features of palatalization and tenseness that applied to all consonants in the language (except *h*). Modern Scottish Gaelic has lost the surface feature of palatalization; many dialects have drastically reduced the range of distinction in the laterals, as well as in other sounds. But in the dialects of Islay (Southern Argyll) I have found 5 distinctive laterals (and 5 nasals). Initially and finally in words there are only 4 distinctions (not directly reflecting in all cases the parent Old Irish state of affairs), but medially we find:

[mʷ ^o lax]	[mulax]	‘top’
[kø ^o lax]	[kolax]	‘rooster’
[ka ^o λax]	[kaλax]	‘old woman’
[ǰv:an]	[ǰvLan]	‘carrying’
[kaλ:er]	[kaL'er]	‘candlestick’

These articulations are sometimes hard to hear, and it is particularly hard to match them naturally and correctly with the initial and final articulation ranges, which involve additional redundant features of affrication, tongue position, devoicing, etc., not found in medial position. But, unusual as the phonetics may be, the systematic relations are clear and, as it turns out, simple. The occurrence of glottalization here matches the phenomenon found throughout these dialects with consonants which we call non-tense when they are in intervocalic position. The range of palatality matches what we note elsewhere in the stops and nasals. Therefore we have, in summary:

	Velar	Palatal	Neither
Lax	<i>l</i>	<i>λ</i>	<i>l</i>
Tense	<i>L</i>	<i>L'</i>	

Unusually rich as these laterals are, they require for their specification nothing more than the system already requires. In fact, Islay seems not to include a separate feature of laterality as we find in the Caucasus and Northwest Coast cases. In Islay the laterals go with, and oppose symmetrically, the nasals.

DISCUSSION

Birnbaum:

If for some languages, such as the Caucasian languages, we can and, indeed, ought to set up a separate feature of laterality, could you give us some idea how such a feature should be defined in articulatory and/or acoustic terms?

Ad Birnbaum: In articulatory terms, all laterals which I have heard can be described more or less as they traditionally have been: By intensive approximation or contact between a mobile portion of the tongue (tip to root) and a region of the roof of the mouth (teeth to velum) such that explosion, friction, or harmonically resonant airstream finds its principal egress through at least one side. In this sense the abstract phonological feature I have discussed matches some simple articulatory parameter of production, and in this sense of perception. I have not yet had an opportunity to study spectrograms of these complex lateral systems. On various grounds I would not expect to find any single simple acoustic parameter common to all lateral instances, but I would be happy to be found wrong.

MUSISCHE FAKTOREN IN DER SPRACHE VON TAUBEN KINDERN

LEON HANDZEL*

Mit der Sprache als Verständigungsmittel befaßt sich u. a. auch die Medizin. insbesondere die Phoniatrie, die vor allem die akustische Struktur als Grundlage für die Bildung der musischen Faktoren der gesprochenen Sprache — Melodie, Dynamik, Sprechrhythmus — untersucht. Diese Faktoren hängen beim Individuum von dessen Begabung, Aufnahmefähigkeit, Temperament, dem vermittelten Inhalt und dem ihn auslösenden Gefühl ab, wofür die Tatsache spricht, daß jeder Mensch das geschriebene Wort auf andere Weise reproduziert und ihm andere Eigentümlichkeiten verleiht. Die Bildung und Gestaltung dieser Faktoren wurde bei 9—13jährigen Kindern untersucht, die verschiedengradige Gehörsschäden aufwiesen, welche zu verschiedener Zeit einsetzten. Das Untersuchungsmaterial bildeten 4 Gruppen von je 30 Kindern mit 1. normalem Gehör, 2. Schwerhörigkeit, 3. Gehörresten, 4. Taubheit. Nach einer Prüfung des Gehörs mit dem Ton-Audiometer der Firma Maico, weiter durch Flüstern, lautes Sprechen und Stimmgabeln, die Romberg-Probe, Gang bei geschlossenen und geöffneten Augen, durch Zeigen und Treffen, erfolgte die Einreihung in die jeweilige Gruppe genau nach dem Audiogramm. Bei Schwerhörigkeit und bei Kindern mit Gehörresten wurde neben dem Audiogramm auch die Verständigungsfähigkeit mit der Umgebung berücksichtigt. Daher sind in einigen Fällen in die Gruppe der Schwerhörigen solche Kinder aufgenommen worden, die trotz ihres audiometrisch festgestellten schlechteren Gehörs im Vorschulalter ihrer Sprache in normalen Verhältnissen entwickelten und sich durch die Sprache mit der Umgebung verständigten, sowie auch Kinder, die später taub geworden sind.

Die Kinder sprachen dreimal den Frage- und Antwortsatz: „Budeš zítra doma?“ und „Budu zítra doma.“ (Bist du morgen zu Hause? — Ich bin morgen zu Hause.) Melodie, Dynamik und Rhythmus, d. h. Tonhöhe, relative Intensität, Silben- und Pausendauer dieser Sätze wurden mit Hilfe des Melodieschreibers von Grützmacher — Lottermoser (modifiziert durch Dr. P. Janota) oszillographisch und sonographisch registriert. Die Analyse dieser Sätze und ihre statistische Bearbeitung ergaben folgende Untersuchungsergebnisse:

* Aus der Phoniatriischen Abteilung der Otolaryngologischen Klinik der medizinischen Akademie in Wroclaw. Leiter der Phoniatriischen Abteilung: Doz. Dr. med. Leon Handzel.

1. Monotonie, Monodynamik und Monorhythmik werden verursacht durch die Wiederholung der Melodie-, Dynamik- und Rhythmusbilder im Bereich jedes Wortes, ja sogar jeder Silbe. Je früher die Gehörsschädigung einsetzt und je größer diese ist, desto deutlicher rücken diese Erscheinungen in den Vordergrund.
2. Infolge der nicht ausgearbeiteten kinästhetisch-verbale Stereotypen sind Artikulationsschwierigkeiten sekundäre Erscheinungen.
3. Je früher sich die Gehörsschädigung einstellte und je größer sie ist, desto mehr erinnern die musischen Faktoren an eine der früheren und niedrigen Formen der ontogenetischen Entwicklung.

Setzte die Gehörsschädigung bei tauben Kindern im frühen Alter ein, so wiesen sie einen solchen Phonations- und später auch Artikulationscharakter auf, wie er für normalhörige Säuglinge und Kleinkinder im ähnlichen Alter typisch ist. Nach den Untersuchungen von Sedláčková handelt es sich hier um

- a) eine Gestaltungskorrelation der Höhe, Intensität und Stimmfarbe,
- b) steigende und fallende Tonhöhenbewegung (in den meisten Fällen),
- c) verlangsamtes Phonations- und später auch Artikulationstempo, was ihre Unterschiedslosigkeit bedeutet und mit der höheren und niedrigeren Lage desselben zusammenhängt. Nicht selten fehlen niedrigere oder höhere Resonanzen und Formanten.

Je größer die Gehörsschädigung ist und je früher sie einsetzt, desto höher ist die Stimme. Dieses Merkmal ist auch Ausdruck der früheren ontologischen Phonations-Entwicklungsstufe.

Die obigen Beobachtungen gestatten folgende Schlußfolgerungen:

- a) Mit der Rehabilitation der musischen Faktoren der Sprache sollte sofort bei der Feststellung der Gehörsschädigung besonders bei resthörigen Kindern begonnen werden.
- b) Unerläßlich notwendig erscheint die Bearbeitung wirksamer Rehabilitationsmethoden von musischen Faktoren besonders für Vorschulkinder.
- c) Die bisherigen objektiven hördiagnostischen Methoden sind durch eine Analyse der Gestaltung musischer Elemente zu ergänzen, wodurch der Hörschädigungsgrad und seine Entstehungszeit präziser beurteilt werden können.

DISCUSSION

Seeman:

Die musischen Faktoren der Sprache haben für ihr Verstehen eine grundlegende Bedeutung. Sie sind der Artikulation übergeordnet. Beim Sprechlernen ahnt das Kind die musischen Faktoren eher nach als die Lautbildung. Ein gestammeltes Wort kann ein hörendes Kind mit völlig veränderten Konsonanten aussprechen und trotzdem versteht jeder seine Bedeutung, da es die Akzente richtig bildet. Die Sprache der Gehörlosen ist unverständlich, weil ihre musischen Faktoren verzerrt sind. Deshalb sollten die Gehörlosenlehrer bei der Sprachbildung ihrer Schüler auf die richtige Entwicklung der Sprachakzente achten.

ÜBER DIE HÖRBARKEIT DER SPRACHMELODIE

LEON HANDZEL und ANNA JAROSZ*

Die Hörbarkeit der Sprachmelodie spielt eine wesentliche Rolle für die Physiologie und Pathologie des Hörens. Das Problem scheint aber in der wissenschaftlichen Literatur noch nicht genügend erörtert zu sein. Dieser Aufsatz stellt einen kurzen Bericht über die Ergebnisse unserer Forschungen dar.

111 Schüler der staatlichen Oberschule in Wroclaw haben dreimal den im Tonband registrierten Satz „Budeš zítra doma“ d. h. „Wirst du morgen zu Hause sein“ gehört. Der Satz wurde als Fragesatz wie auch als Antwortsatz ausgesprochen. Jeder Schüler registrierte in üblichen Notenzeichen die Melodie der beiden Sätze. Die subjektiv erfaßte Tonhöhe der einzelnen Silben (d. h. die Satzmelodie) wurde nachher statistisch bearbeitet, so daß a) der Mittelwert $[M]$, b) der durchschnittliche Fehler des Mittelwertes (etha), c) die durchschnittliche Abweichung (sigma) berechnet wurde. Die Ergebnisse sind in Tabellen (I und II) dargestellt, außerdem sind in den Diagrammen die Mittelwerte in der üblichen Tonleiter veranschaulicht (Diagramm Nr. 1 und 2).



Abb. 1.

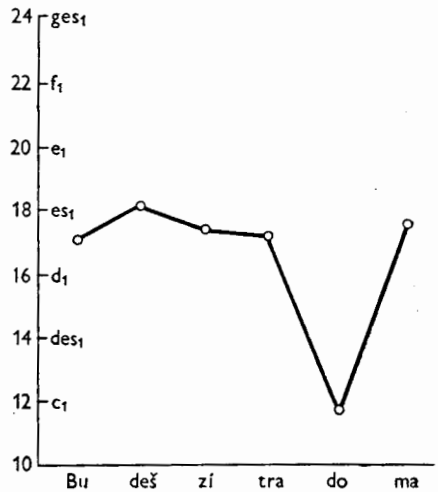


Abb. 2.

* Aus der Phoniatischen Abteilung der Otolaryngologischen Klinik der Medizinischen Akademie in Wroclaw, Leiter der Phoniatischen Abteilung: Doz. Dr. med. Leon Handzel.

Der Mittelwert der Melodie (d. h. die Tonhöhe oder Tonabstand zwischen den einzelnen Silben) des Fragesatzes und des Antwortsatzes gibt die charakteristische Melodie des Frage- beziehungsweise des Antwortsatzes wieder.

Die steigende Intonation, typisch für den Fragesatz, wurde von den Schülern perzipiert und in den Noten registriert. Das ersieht man aus dem wichtigen Höhenunterschied, der mehr als zwei etha zwischen der 4—5 und 5—6 Silbe, beträgt. Der durchschnittliche Unterschied zwischen der höchsten und der tiefsten Silbe schwankt etwa zwischen 3 Halbtönen. Was die Tonleiter betrifft, liegt die durchschnittliche Melodie des Fragesatzes etwa zwischen c 1 — dis 1, der durchschnittlicher Fehler (für 2 etha berechnet) beträgt ein wenig mehr als einen Halbton, daß heißt einen kleinsten musikalischen Intervall. Die Standardabweichung für 1 sigma schwankt zwischen 6—7 Halbtönen.

Die fallende Intonation, die charakteristisch für den Antwortsatz ist, drückt sich in der Satzmelodie aus. Für die Melodie ist ein unwichtiger Höhenunterschied zwischen der 1.—2. und der 3.—4. Silbe charakteristisch, der zwischen der 2. und 3. Silbe sicherlich mehr als 2 etha beträgt. Außerdem ist der Höhenunterschied etwas größer als 2 etha. Der Höhenunterschied zwischen der höchsten und der tiefsten Silbe schwankt etwa zwischen 6—7 Halbtönen. In der üblichen Tonleiter registriert, ist die Melodie des Antwortsatzes zwischen h-1 und ges-1 enthalten.

Der durchschnittliche Fehler des Mittelwertes, für 2 etha berechnet, schwankt in den Grenzen des kleinsten musikalischen Intervalls. Die Standardabweichung für 1 sigma beträgt etwa 5—6 Halbtöne. Der durchschnittliche Fehler des Mittelwertes (also auch die durchschnittliche Abweichung) ist in einem Fragesatz etwa höher als in einem Antwortsatz. Die erste Silbe des Antwortsatzes hat den geringsten durchschnittlichen Fehler des Mittelwertes und die geringste durchschnittliche Abweichung, die nicht mehr als einen Halbton beträgt. Die zweite Silbe hat eine größere Abweichung. Die zweite Silbe hat auch den höchsten Ton; sie ist also die höchstliegende innerhalb übrigen Silben.

Die statistische Analyse hat ergeben, daß die Schüler die einzelnen Tonhöhen verschieden perzipieren und registrieren, und zwar etwa zwischen 5—7 Halbtönen für 1 sigma. Der durchschnittliche Fehler des Mittelwertes ist trotz der großen Quote der untersuchten Schüler verhältnismäßig gering. Er beträgt nämlich etwa 1 Halbton. Ähnliche Durchschnittszahlen (der Durchschnittsfehler und die Standardabweichung) für subjektive Bewertung der Silbentonhöhe haben auch Sedláček und Sychra, die die relativ subjektive Hörbarkeit der Tonsilbenhöhe bei Menschen mit relativem musikalischen Gehör prüften, errechnet.

Wenn wir die Ungleichartigkeit der subjektiven Hörbarkeit der Sprachmelodie (d. h. das Hören der Silbenhöhe) betrachten, so sehen wir, daß dabei verschiedene Faktoren eine wichtige Rolle spielen. Man kann annehmen, daß bei der Anerkennung der Tonsilbenhöhe die Qualität des menschlichen Gehörs und die physiologisch-psychologischen Eigenschaften und Eigentümlichkeiten des Hörens äußerst wichtig sind.

Vergleichend das subjektive Hören mit dem objektiven Standardtest der Satzmelodie, stellen wir fest, daß je tiefere (bzw. höhere) Silbentöne auftreten und je größer die Tonhöhenbewegungen erscheinen, desto größere Unterschiede in der subjektiven Perzeption (und auch in der subjektiven Notierung) sind. Diese Beobachtung stimmt mit den Untersuchungen von Sedláček und Sychra überein. Das menschliche Gehör ist ontogenetisch bedingt; diese Feststellung geht aus der Tatsache hervor, daß je mehr irgendein Ton von dem Ton a-1 entfernt ist, desto schwieriger der Mensch die Tonhöhe erkennt.

Bei der Schätzung des musikalischen Gehörs, besonders bei Menschen, die sich mit musikalischem Folklore professionel beschäftigen oder beschäftigen wollen, ist die Untersuchung ihrer musikalischen Gehörs auf Grund der Sprachmelodie besonders wichtig. Das richtige Hören ist für die Autentität der Melodie der Volkslieder entscheidend. Wenn dies nicht der Fall ist, können verschiedene Melodie- oder Liedervariationen notiert werden, die in Wirklichkeit nicht existieren.

Tabelle 1. (Fragesatz)

Silbe	M	etha	sigma
bu	16,95	0,61	6,50
deš	18,04	0,64	6,81
zí	17,31	0,64	6,83
tra	17,13	0,63	6,63
do	11,60	0,63	6,72
ma	17,50	0,64	6,84

Tabelle 2. (Antwortsatz)

Silbe	M	etha	sigma
bu	23,48	0,44	4,73
du	22,73	0,52	5,51
zí	17,17	0,58	6,11
tra	16,04	0,55	5,80
do	13,23	0,55	5,81
ma	10,62	0,54	5,72

DISCUSSION

Stock:

Wir beobachteten, daß die Tonhöhe betonter Silben „genauer“ gehört wird. Zeigt sich das auch bei Ihnen, ist vielleicht die Streuung der Notenwerte hier geringer als in den Urteilen

über unbetonte Silben? Wir beobachteten ferner, daß man 2 Typen von Hörern unterscheiden kann; manche hören betonte Silben immer höher und andere hören trotz Akzentuierung rel. genau die Silbmelodie. Haben Sie ähnliche Erfahrungen gemacht?

Handzel:

ad Stock: Die Beobachtungen von Dr. E. Stock können wir, auf Grund unserer Forschungen bestätigen. Wegen der Kürze des Vortrages konnten wir sie nicht besprechen.

INTERCHANGE IN SPEECH COMMUNICATION

L. S. HARMS*

1.0 INTRODUCTION

1.1. The title of this paper expands to read: interchange in a two-person speech-communication system. The term system retains its usual meaning of organized, task-oriented activity within a structural framework. Communication systems indicate the general area of interest, and a speech-communication system designates the particular area of interest. Two-person specifies the basic unit. Interchange names the particular pattern of activity within a two-person speech-communication system of concern in this paper.

1.2. A model of a two-person system is shown in Figure 1. Notice there are two-persons joined in a particular speech space to constitute a communication system. The receptors correspond to concerns in auditory phonetics, the brain to neural phonetics, the articulators to articulatory phonetics, and speech space to acoustic phonetics.

1.3. In such a two-person system, the capacity of any one sub-system affects the capacity of the entire system. Thus if Person B has a mild hearing loss, the capacity of the entire system is affected. If Person A has a mild hearing loss, and a limited amount of time, and Person B is speaking in a second language, and the subject is political on which B and A strongly disagree, the probability of a productive outcome approaches zero.

2.0 STRUCTURE

In a two-person system, three basic structural patterns are evident. These patterns generate from two social control arrangements and three message sequences. These basic structures are called: interview, tutorial and interchange.

2.1. In the *interview*, the interviewer controls. He asks directive questions. The interviewee answers these questions. Tasks which can be accomplished are limited

* Speech Communication Center, University of Hawaii, Honolulu, USA.

in number, and outcomes are largely predetermined. A question constrained message is generated.

2.2. In the *tutorial*, the tutor controls. He makes informative statements. The student asks questions. Tasks are diverse, and the range of outcomes are predictable. A branched message is generated.

2.3. In the *interchange*, the two-persons share control. Both persons ask questions and make statements. Tasks are highly diverse, and outcomes are not readily predicted. An adaptive message is generated.

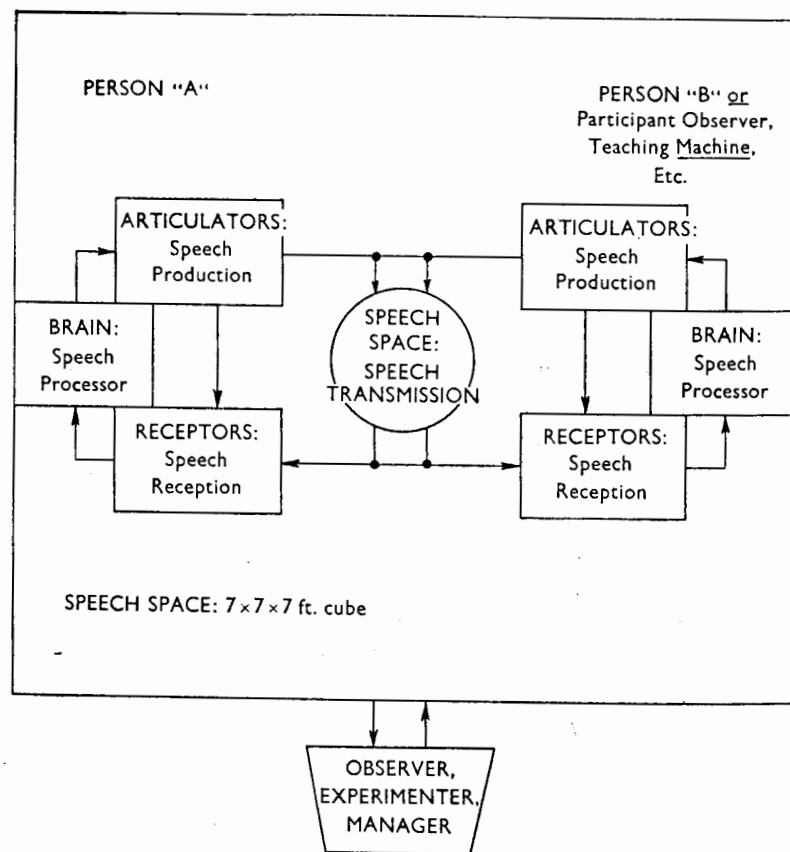


Fig. 1.

3.0 SOME PROBLEMS IN INTERCHANGE

In this section, I wish to focus only on the speech communication structure I have called interchange. Further, I wish to continue to limit the discussion to a two-person system. Of particular interest are two-persons of substantially different linguistic-cultural backgrounds engaged in task-oriented speech communication.

3.1. In the following illustration, two persons said word lists to each other. Person A said a word; Person B repeated it. Then, Person B said a word and Person A repeated it. And so on. The monosyllabic word lists contained a sample of syllable types and phonemes of mid-west American English. The persons were paired under the following conditions:

Language Background	Person A	Person B
1 Same	Native speaker	Native speaker
2 Same	Non-native speaker	Non-native speaker
3 Different	Native speaker	Non-native speaker
4 Different	Non-native speaker	Non-native speaker

At a distance of three feet in a sound treated room, the accuracy of repetition under the four conditions varied in the following manner. Under the condition of the same language background, two persons achieved a low and stable error rate. Interestingly, the native language, whether English, French, or Japanese, did not appear to affect the error rate on the English word lists. This finding holds only for two persons with the same language background.

Under the conditions of different language background, two persons of dissimilar native language background achieve a high error rate. As long as the native language was different a high error rate was obtained. Thus, on the English word lists, a native French speaker and a native Japanese speaker scored similarly, for instance, to a native English speaker and a native French speaker.

Errors were scored on degree of phoneme correspondence between a monosyllabic word said by one person and repeated by the other. The persons with different native language backgrounds achieved approximately three times as many errors as persons with similar language backgrounds. This finding demonstrates that a native speaker of Y finds another native speaker of Y more intelligible in language X than he does a native speaker of language X.

At this point, one may question the manner of error classification. A prescriptive or standard language target suggests one solution; a communication system approach suggests another. Within two-person interchange, the decision the native speaker makes in this regard directs him either to *adapt* to the non-native speaker, or to delay task completion until he can train the non-native speaker to adequately approximate his own speech, or to abandon the attempt to communicate with the non-native speaker.

3.2. A second type of illustration, which depends on the findings of the first, concerns the structure brought about by three types of sentence by sentence social control. These are: the interview, in which the interviewer controls by continuing to pose questions; the tutorial in which the tutor controls by making task-oriented statements; and, the interchange in which both persons share control, that is, both freely ask questions, and make statements.

Two persons are assembled in a quiet room. They are instructed to complete a task,

or arrive at an agreement on an assigned topic. Tasks and topics are selected to require sustained interchange, that is a free exchange of questions and statements.

Under these conditions, as two persons differ in linguistic-cultural background, an interesting transformation occurs. One or the other, usually the native speaker, assumes control. He either conducts an interview, or offers a tutorial. That is, he controls by asking questions in the manner of an interviewer, or he controls by making statements in the manner of a tutor.

The previous finding cited on the accuracy of word reception and other findings, indicate that as persons differ in background, frequent checking procedures are required if speech-communication is to be sustained for that amount of time required to complete a task or achieve a working agreement.

The interchange, or shared-control structure, appears best suited to effective and satisfactory outcomes. However, as difference in background between two persons increases, the probability that one speaker—usually the native speaker—will control through questions or statements also increases. As this control transformation occurs the probability of a mutually satisfactory outcome decreases.

3.3. Within a two-person system, the degree of correspondence between the phonemic and syntactic patterns of the two-persons tends to produce control structures which limit both the range of tasks and topics of speech-communication, and the probability of satisfactory outcomes. What is interesting is that the native speaker who assumes control is likely to perceive himself as helpful while the second language speaker is likely to perceive him as aggressive.

4.0 SUMMARY

My paper today has focused on sustained communication between two persons of dissimilar linguistic-cultural backgrounds. From a system view, I have attempted to show that degree of correspondence between speech patterns of the two persons affects what can be said which in turn limits the range of outcomes.

DISCUSSION

Black:

Are the interesting results to be generalized over a face-to face situation *and* a remote "Telephone" system?

Marquardt:

Professor Harms' excellent presentation suggests that a new emphasis in second-language-teaching should be given to a possibly new discipline which might be called cross-culture-communication. This new discipline would focus attention upon kinds of adjustment communicators must make in cross-culture interaction and reduce emphasis upon such exercises in the language class as pattern-practice and pronunciation drill. I would like to ask whether Prof. H. has thought out ways in which the principles he has discoursed might be extended to class-room activity or to the development of teaching materials.

EXPERIMENTS WITH ARTIFICIAL INTONATION CONTOURS

J. 'T HART—A. COHEN*

In analysing perceptually Dutch pitch patterns, use is made of the Intonator,^{1,2} the various applications of which are to be reported here. These can be viewed from three aspects, leading to experiments: a) in an effort to check hypotheses derived from previous analysis, b) involving a technique of analysis by synthesis to uncover new phenomena whose relevance, as well as the perceptual tolerances of already known patterns, are to be investigated, and c) calculated to afford insight into a speaker's assumed capacity, on the basis of a deliberate programme, to control the up and down movements in vocal pitch.

a) As for the "declination", the often observable gradual fall of pitch during an utterance, it was found to be generally applicable in tests in which short utterances (2—3 sec), processed through the Intonator, were presented to subjects. With utterances exceeding 5 sec, the rule controlling the slope had to be modified to provide for a levelling off towards the end, a result, later confirmed by the observation of a natural lower limit in recordings of voice pitch. The intuitive notion of dominance, used as a concept to determine the locations of rises and falls (constituting a "hat pattern"²), could be confirmed in a reading experiment in which subjects were asked to underline in a continuous text those words that, in their opinion, were dominant.

A systematic alternation of rises and falls, implied in the basic shape of the hat pattern, was found to allow for a more explicit formulation, viz. "what goes up must come down", thus precluding a succession of, e.g., two rises. However, in some cases a succession of falls was observed; use of the Intonator revealed, that in such cases the second fall was optional. Such a test with the Intonator in deciding which of the two semifalls is perceptually more relevant affords a typical example of the method of analysis by synthesis.

b) This very method was used to establish perceptual tolerances of locations in time of major features of the hat pattern. The values found could amount to as much

* IPO, Eindhoven, Holland.

¹ A. Cohen, J.'t Hart; Perceptual analysis of intonation patterns, *Proc. 5th ICA*, Liège, 1965.

² J.'t Hart; Perceptual Analysis of Dutch Intonation Features, *IPO Annual Progress Report 1*, 1966.

as 100 ms or more, unless an adjacent syllable would adopt spurious prominence, in which case a change of some 30 ms could bring about a drastic change in interpretation of the utterance.

Utterances containing odd numbers of dominant words will show a succession of rises, seemingly contradicting the rule mentioned above. To uphold the rule,

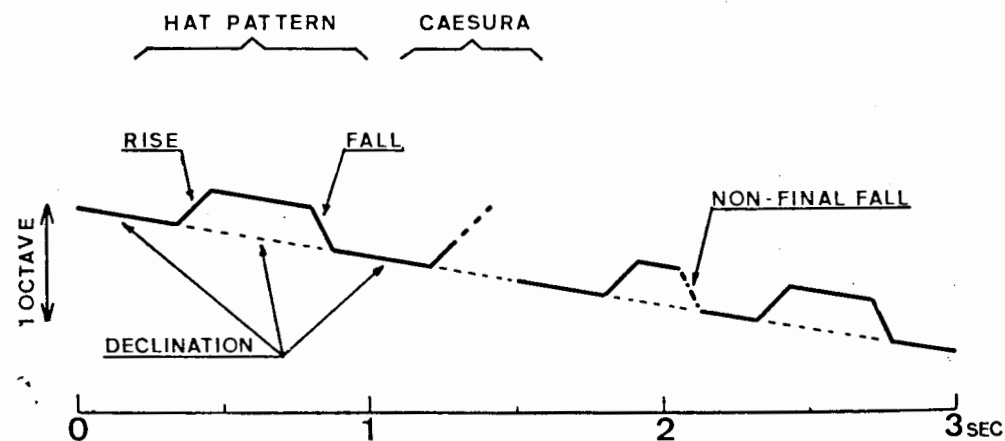


Fig. 1. Illustration of some of the most important features of Dutch intonation, in stylized form, and for an imaginary utterance.

a fall had to be introduced in between two rises, without causing any additional prominence. Such a "non-final" fall, devised on the basis of the model, and successfully tested for its perceptual relevance, proved to be helpful in unravelling the seemingly capricious up and down movements found in pitch recordings.

In a study of Dutch proverbs the notion of caesura (a sign of continuation between two hat patterns) could be verified by simulating pitch patterns to obtain best matches for naturalness with those of the spoken versions.³ It was found, that it had to be considered more complex than originally assumed, and capable of influencing the preceding hat pattern. This anticipatory action involves either an omission of the final fall preceding the caesura, or its substitution by a rise, requiring the introduction of a non-final fall to obviate an otherwise forbidden succession of rises. Of the two main characteristics of a caesura, the upward wriggle has disappeared, whereas resumption of the pitch at a low level has been preserved. The modification of the caesura rules helped in analysing pitch curves while studying observed errors occurring under conditions of reading aloud, as will be dealt with under c).

c) Material, obtained from having subjects read aloud texts presented piecemeal and devoid of punctuation, provided errors of intonation. Such errors, which were

detected on an intuitive basis, could be interpreted in terms of a possible disarray on the part of the speakers, confronted as they are with insufficient data on which to base an adequate control of up and down movements of voice pitch. Typical errors are inconsistencies in generating permissible variants of a caesura and executions of final falls on non-dominant terminal words of stretches that semantically require an ending. Similar errors can frequently be observed in the speech of news-readers. Processing samples of these errors by means of the Intonator and based on rules implied in the model led to significant improvement.

Such a study of errors may help in getting a better insight into the mechanism underlying the correct programming of voice pitch, and reflecting in some way or another the syntactic structure in normal speech.

³ J. F. Rijekaert, *IPO Rapport no. 89* (In Dutch.).

LAUTDAUER ALS MERKMAL DER WAHRGENOMMENEN QUANTITÄT, QUALITÄT UND BETONUNG IM DEUTSCHEN

GEORG HEIKE

In diesem Beitrag soll über Teilergebnisse linguistischer Hörtests berichtet werden, in denen ausgewählte Testbeispiele mit systematisch variiert Vokaldauer benutzt wurden. Die Ergebnisse werfen Fragen auf, zu deren Beantwortung sich weitere Untersuchungen als notwendig erweisen. Als Ausgangsmaterial für den ersten Hörtest, über den berichtet wird, dienten die vom Verfasser gesprochenen Wörter „Aas“ („aß“), „Aale“ („Ahle“) und „Ohm“. Mit Hilfe eines Segmentators und

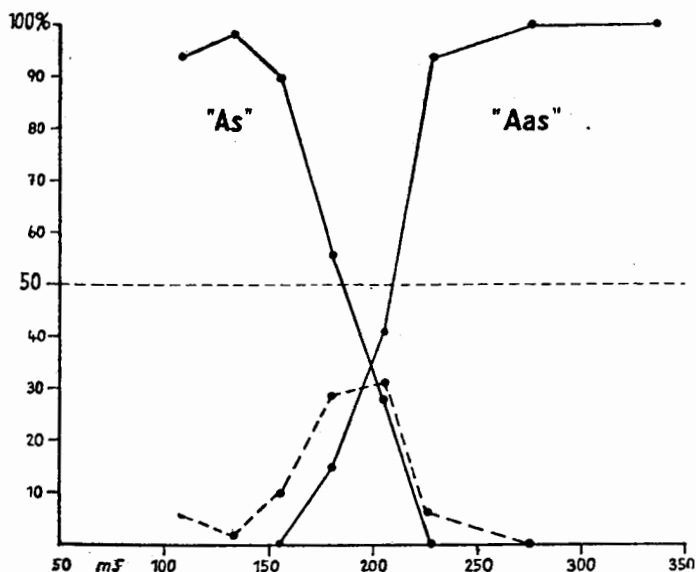


Bild 1: Antworten in % "eindeutig 'Aas' bzw. 'As' ",
----- : "nicht der Norm entsprechend".

Diagramm I.

oszillographischer Kontrolle wurden die Vokale initial um Beträge abnehmender Dauer verkürzt. Jedes der resultierenden Testbeispiele wurde viermal kopiert und mit für eine andere Fragestellung vorgesehenen Wörtern in zufälliger Reihenfolge einer Gruppe von 12 deutschsprachigen Hörern dargeboten. Die Hörer sollten das

Wort notieren, das sie eindeutig verstanden hatten. Eine als „nicht der Norm entsprechend“ empfundene Aussprache sollte gesondert bezeichnet werden. Aus den Diagrammen nach Bild 1 bis 3 können folgende allgemeine Schlüsse gezogen werden: (1) Nach initialem Verkürzen von Vokalen mit sog. ‚phonologischer Länge‘, wobei der Übergang Vokal-Konsonant unverändert bleibt, werden von naiven Hörern

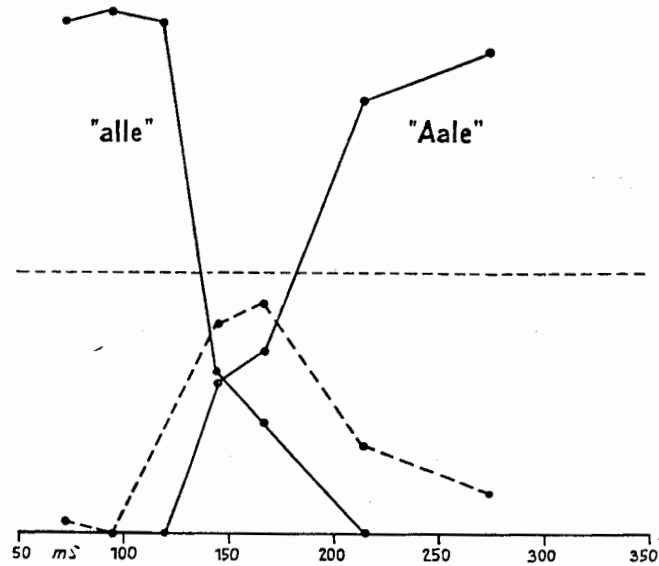


Bild 2: Antworten in % „eindeutig ‚Aale‘ bzw. ‚alle‘“,
-----: „nicht der Norm entsprechend“.

Diagramm 2.

in einem bestimmten Dauerbereich in mehr als 80 % der Fälle Wörter mit Vokalen sog. ‚phonologischer Kürze‘ gehört. 80 % der Urteile ‚Kürze‘ (d. h. genauer: ‚As‘, ‚alle‘, ‚un‘) wurden bei einer Vokaldauer $T < 160$ ms und ‚Länge‘ (d. h. genauer: ‚Aas‘, ‚Aale‘, ‚Ohm‘) bei einer Dauer $T > 250$ ms abgegeben. Diese Ergebnisse deuten darauf hin, daß der zu beobachtende unterschiedliche Intensitätsverlauf des finalen Vokalteils, der in Zusammenhang mit dem Gegensatz ‚fester‘ und ‚loser Anschluß‘ gebracht wird,¹ im Vergleich zur Vokaldauer unter perzeptivem Aspekt irrelevant ist. (2) Im Falle von Vokalen mit mittlerem Kompaktheitsgrad findet zusätzlich eine Änderung der wahrgenommenen Vokalqualität statt (Bild 3). Dieser Sachverhalt wurde entsprechend bei [e:] (\rightarrow °[I]) und [ɸ:] (\rightarrow °[Y]) festgestellt, worüber hier kein Material vorgelegt wird. (3) Die Häufigkeitsverteilung der Antworten „nicht der Norm entsprechend“ deckt sich mit dem Übergangsbereich

¹ Essen, O. von: Trubetzkoy's „fester“ und „loser Anschluß“ in experimentalphonetischer Sicht. *Proc. 4th Int. Congr. Phonet. Sciences*, The Hague (1962), 590—597.

zwischen Antworten ‚Kürze‘ und ‚Länge‘. — Folgende Differenzen zwischen den drei Diagrammen können beschrieben werden: (1) Der Kreuzungspunkt der Urteilskurven und die Mitte des Unsicherheitsbereiches (zwischen 50 % ‚Kürze‘ und ‚Länge‘) zeigen bei Diagramm 2 niedrigste Dauerwerte. Die Lage der Urteilskurve für ‚Kürze‘ zu niedrigeren Dauerwerten hin hängt vermutlich mit der Zweisilbigkeit zusammen.

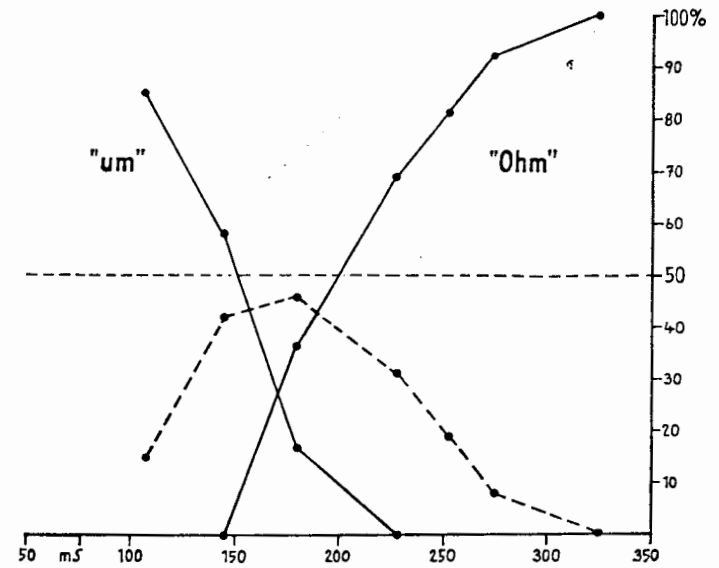


Bild 3: Antworten in % „eindeutig ‚Ohm‘ bzw. ‚um‘“,
-----: „nicht der Norm entsprechend“.

Diagramm 3.

(2) Hinsichtlich der Größe des Unsicherheitsbereiches, sowie der Höhe und Erstreckung der Urteilskurven „nicht der Norm entsprechend“ weist Diagramm 3 die größten Werte auf, was offensichtlich damit zusammenhängt, daß hier zusätzlich die Wahrnehmung der Vokalqualität sich ändert. Ebenso ist die Steilheit der Urteilskurve für „um“ am geringsten. (3) In Diagramm 1 steht ein schmaler Unsicherheitsbereich (ca. 25 ms) im Zusammenhang mit einer bemerkenswerten Urteilssicherheit und -konstanz über einen weiten Dauerbereich. Es kann deshalb prinzipiell auf ein ausgeprägtes Unterscheidungsvermögen für absolute Dauerwerte mit linguistischer Funktion geschlossen werden. — Diese Untersuchungen sollen weitergeführt werden, wobei unter anderem die Rolle der Folgekonsonanten, der Silbenzahl und des Sprechtempos eines gesprochenen Kontextsatzes für die Beurteilung zu prüfen ist. Ferner wäre von Interesse festzustellen, ob im Übergangsbereich zwischen wahrgenommener ‚Kürze‘ und ‚Länge‘ die Diskrimination von Dauerunterschieden sinnfreier Signale größer ist, als in den Bereichen eindeutiger Zuordnung.

Bild 4 zeigt die Ergebnisse eines Hörtests, in dem 10 Hörer aufgefordert wurden, den als „betont“ gehörten Vokal sinnfreier Testbeispiele anzugeben. Die Beispiele

wurden aus zwei gehalten gesprochenen [a]-Vokalen verschiedener Grundfrequenz (110 Hz und 120 Hz) derart hergestellt, daß das Intensitätsverhältnis zweier durch eine Pause von 180 ms getrennter Vokale systematisch variiert wurde. Zwei Vokaldauern fanden Verwendung: 70 ms und 210 ms. Von den vielfältigen möglichen Interpretationen der Diagramme seien die folgenden genannt: (1) Der zu erwartende

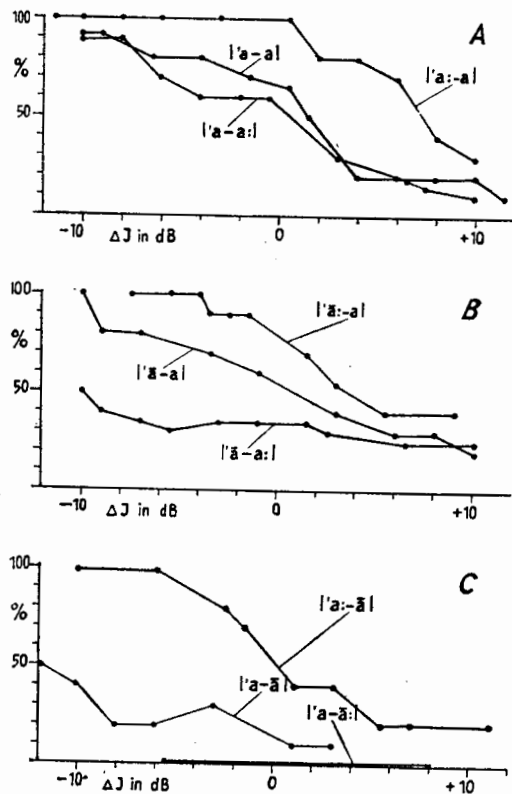


Bild 4: Antworten in % "erster Vokal betont"
Diagramm 4.

Sachverhalt, daß größere Dauer mit häufigeren Betonungsurteilen korreliert, wird bestätigt. (2) Bei gleicher Tonhöhe, Intensität und Dauer der beiden Vokale überwiegt wahrgenommene Erstbetonung, die erst durch größere Dauer des 2. Vokals auf ca. 50 % Urteile kompensiert wird (A). (3) Niedrigere Tonhöhe des 2. Vokals führt zu einer unerwarteten Verminderung der Urteile „Erstbetonung“ (B). (4) Höhere Tonhöhe des 2. Vokals resultiert in einer weiteren Erhöhung der Urteile „2. Vokal betont“ (C). Verbunden mit größerer Dauer des 2. Vokals sind im ganzen untersuchten Bereich der Intensitätsdifferenzen 100 % Urteile „2. Vokal betont“ abgegeben worden. (5) Zum Verhältnis von Tonhöhe und Dauer läßt sich nur bspw. feststellen, daß ein Grundfrequenzverhältnis von 110 zu 120 Hz gleichlanger Vokale (C) zu häufigeren Urteilen „2. Vokal betont“ führt, als ein Dauerverhältnis

von 70 zu 210 ms bei gleicher Grundfrequenz. — Diese Ergebnisse sollten durch weitere Tests vervollständigt werden und als Grundlage für Tests mit sinnvollem Material dienen.

DISCUSSION

Maiwald:

Sie haben im zweiten Teil Ihres Vortrages von Untersuchungen berichtet über die Verschiebung der Betonung in zwei aufeinanderfolgenden Vokalen durch Änderung der Dauer, Tonhöhe, beziehungsweise Lautstärke eines der Vokale. Die geschilderte Untersuchungsmethode scheint mir geeignet, wesentliche Ergebnisse im Sinne prosodischer Regeln zu liefern. Können Sie aufgrund Ihrer bisherigen Versuche schon quantitative Angaben über den wechselseitigen Einfluß dieser Parameter auf die Betonung oder vielleicht über ihre Wertigkeit machen?

Martens:

Vorschlag: Wenn man einen gemeinsamen Nenner finden will für die Tatsache, daß gekürzte [o] als [u] und gekürzte [a] als [ɑ] aufgefaßt werden, bietet sich an, dies als einen Übergang aus der Kategorie „gespannt“ in die Kategorie „ungespannt“ zu bezeichnen. Damit hätte der gleichartige Vorgang eine gleichartige Bezeichnung (wenngleich dadurch nur der gleiche Teil bezeichnet ist, denn im Grunde ist die Deutung von kurzem [o] als [v] etwas anderes als die Deutung von kurzem [a] als [ɑ]. Aber Heike bestätigt die Methode der Sprachpädagogen, ein offenes ungespanntes [u] aus [o] oder [ɪ] aus [e] und eine [ʏ] aus [ø] zu gewinnen!)

Richter:

Die Experimente des Referenten unterstreichen einmal mehr, daß die signalmäßige Differenzierung für die Identifikation lautlicher Einheiten eine geringere Rolle spielt, als der Vorstellung entsprechen würde, daß das Sprachsignal in einfacher Weise phonemisch segmentiert werden kann.

In weiteren Versuchen sollten die Sprecher systematisch nach der regionalen Herkunft variiert werden, da sich die deutschen „Regionallautungen“ danach unterscheiden dürften, welcher der Faktoren „Dauer“, „Lautheit“ oder „Melodie“ bei der Determination der Betontheit einer Silbe führend ist.

Heike:

Die „Wichtigkeit“ der untersuchten Parameter ist schwer abschätzbar, weil sie verschiedenen unvergleichbaren Wahrnehmungsqualitäten entsprechen. Es ließen sich höchstens nach Durchführung entsprechender Tests, wie sie zur Mel- und Soneskala geführt haben, Angaben über die unterschiedliche Zahl von Valenzen machen, die bei verschiedenen Parametern die gleiche Urteils-häufigkeit ergeben.

Heike:

ad Maiwald: Quantitative Angaben über die Größe jedes Parameterwerts, bei dem unter Konstanzhaltung der anderen, eine bestimmte Urteils-häufigkeit abgegeben wird, sind aus den Diagrammen zu entnehmen. [Weitere Einzelheiten siehe meine Arbeit 'Suprasegmentale Analyse'. Marburg (demnächst bei Elwert, 1969). Die Wertigkeit der Parameter ließe sich eventuell erst abschätzen, wenn bei den verwendeten Signalen die energetischen Unterscheidungsschwellen der Parameter bestimmt worden sind, wie dies bei Sinusoidalsignalen bereits der Fall ist.

ÜBER FAKTOREN, WELCHE DIE PERZEPTION DER INDIVIDUELLEN HAUPTSPRECHTONHÖHE FÖRDERN ODER HEMMEN

LISELOTTE HERBST*

Die individuelle Hauptprechtonhöhe (HST-Höhe) eines Menschen ist als akustisches Sprachsignal anzusehen. Sie ist einerseits an anatomisch-physiologische Faktoren, andererseits an psychische Vorgänge gebunden. Der Hörer kann an der perzipierten HST-Höhe Alter, Geschlecht und Konstitution, aber auch den emotionalen Zustand eines Sprechers einschätzen. Um Normwerte für die individuellen physiologischen HST-Bereiche zu ermitteln, wurden von 2124 deutschen Versuchspersonen (Vpn) die verwendeten HST-Höhen bei unterschiedlichen Sprechleistungen untersucht, und zwar im Gespräch beim Reihensprechen, beim Summen des stimmhaften S-Lautes, bei der Kauphonation, beim Vorlesen eines ungespannten Prosatextes (Goethe) sowie eines gespannten Poesietextes (Schiller). Die Beobachtungen, die bei der Bestimmung der verwendeten HST-Höhen vermerkt werden konnten, bilden die Grundlage meiner Ausführungen. Bei der experimentell-auditiven Untersuchung des HSTs genügt nicht der Vorgang der Perzeption, sondern hierbei wird die Apperzeption notwendig. Es muß nämlich eine sofortige Wertung aller perzipierten Sprechtonhöhen vorgenommen werden, damit die charakteristische HST-Höhe im Verlauf der auf- und absteigenden Gleittöne apperzipiert werden kann. Schlußfolgernd habe ich in meinen Ausführungen den Terminus Apperzeption verwendet.

Die Abb. 1 zeigt die Lage des apperzipierten HSTs im Verhältnis zum Gesamttonhöhenverlauf des von einer weiblichen Vp gesprochenen Satzes: „Heute ist der 13. März 1964“. Die obere Kurve ist die Intensitätskurve, die darunterliegende die Tonhöhenkurve. Die Lage des HSTs ist mit einer hindurchgezogenen Geraden gekennzeichnet. Es ist ersichtlich, daß sich das Ohr bei der Apperzeption des HSTs vorwiegend an die Laute und Silben halten muß, die mit mittlerer Intensität gesprochen werden und die sich in mittlerer Sprechtonlage befinden. Die Apperzeption der HST-Höhe wird 1. gefördert, wenn der Untersuchende über ein sicheres Unterscheidungsvermögen der Tonhöhen sowie über ein gutes Konzentrationsvermögen verfügt; 2. ist die Apperzeption günstiger, wenn der Sprecher eine gesunde Stimme hat; 3. ist fördernd, wenn der Sprecher die hochdeutsche Satzintonation verwendet; 4. spielt die Größe des Sprechstimmumfangs eine Rolle. Je kleiner er ist, umso

* Pädagogisches Institut Halle (Saale), DDR.

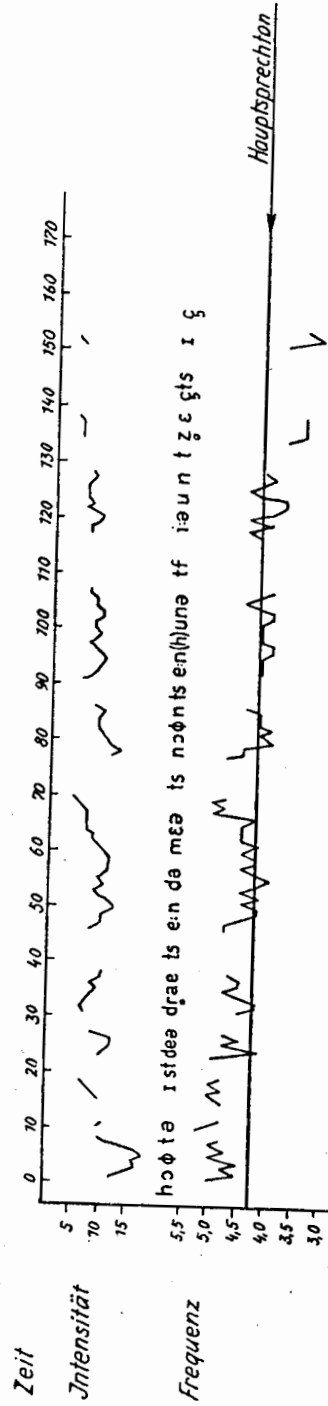


Abb. 1. Tonhöhenverlauf der gesprochenen Sätze (Ansage des Datums) — Frauen
 Zeit: 1 = 20 ms
 Intensität: 1, 2 usw. als Vergleichseinheit
 Frequenz: 5,5 = 284,9 — 255,0 Hz
 5,0 = 254,9 — 235,0 Hz
 4,5 = 234,9 — 205,0 Hz
 4,0 = 204,9 — 185,0 Hz
 3,5 = 184,9 — 155,0 Hz
 3,0 = 154,9 — 135,0 Hz

leichter wird das Erfassen des HSTs; 5. kann der HST bei jenen sprachlichen Leistungen gut apperzipiert werden, bei denen sich anhand von instrumentellen Registrierungen nachweisen läßt, daß der Mittelwert des verwendeten Sprechstimmumfangs mit dem Durchschnittswert der registrierten Frequenzen und der am häufigsten verwendeten Sprechtonhöhe identisch ist, wie es die Abb. 2 zeigt.

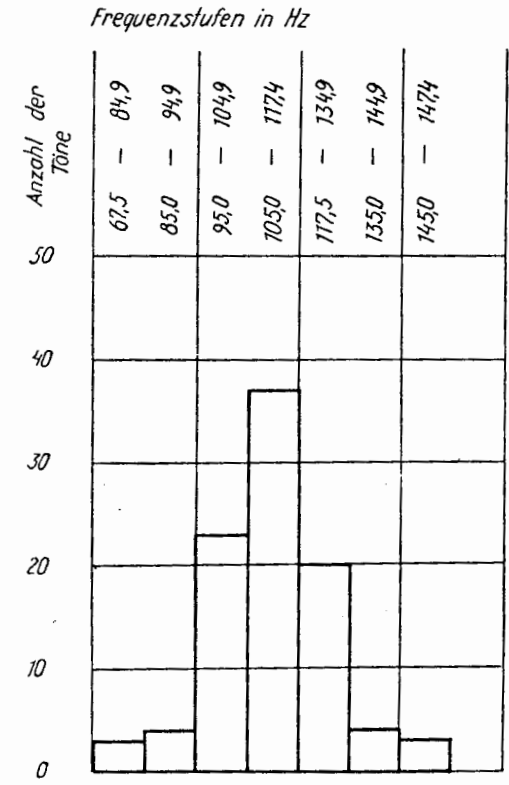


Abb. 2. Darstellung der Häufigkeit der verwendeten Sprechtonhöhen — Männer Satz (Ansage des Datums)

Diesem Säulendiagramm liegt ebenfalls die Ansage des Datums: „Heute ist der 13. März 1964“, von einer männlichen Vp gesprochen, zugrunde. Die Höhe der Säulen zeigt die Häufigkeit der verwendeten Tonhöhen (in Frequenzstufen) an. Es ist festzustellen, daß die registrierten Sprechtonhöhen annähernd symmetrisch um einen mittleren Gipfel angeordnet sind. Die Vp verwendete die hochdeutsche Satzintonation, normales Sprechtempo und natürlichen Satzrhythmus. Die Apperzeption des HSTs war einfach. Die Apperzeption des HSTs wird 1. behindert, wenn eine geringe Hörschulung des Untersuchenden für diese Tätigkeit vorhanden ist und wenn Ermüdungserscheinungen auftreten; 2. wird die Apperzeption erschwert, wenn der Sprecher verhaucht oder belegt spricht, weil hierbei der Grundton meist schwer zu bestimmen ist; 3. wirkt behindernd, wenn der Sprecher eine ungewöhnliche Satzintonation verwendet. Vermutlich spielt das Problem des funktionellen Nach-

vollziehens bei der Apperzeption mit eine wichtige Rolle; 4. ist bei stark affektgeladener Redeweise das Erfassen des HSTs schwierig, weil der Sprechstimmumfang oft große Ausmaße annimmt; 5. ist die Apperzeption des HSTs in den Fällen ungünstig, wo der Mittelwert des verwendeten Sprechstimmumfangs und der Durch-

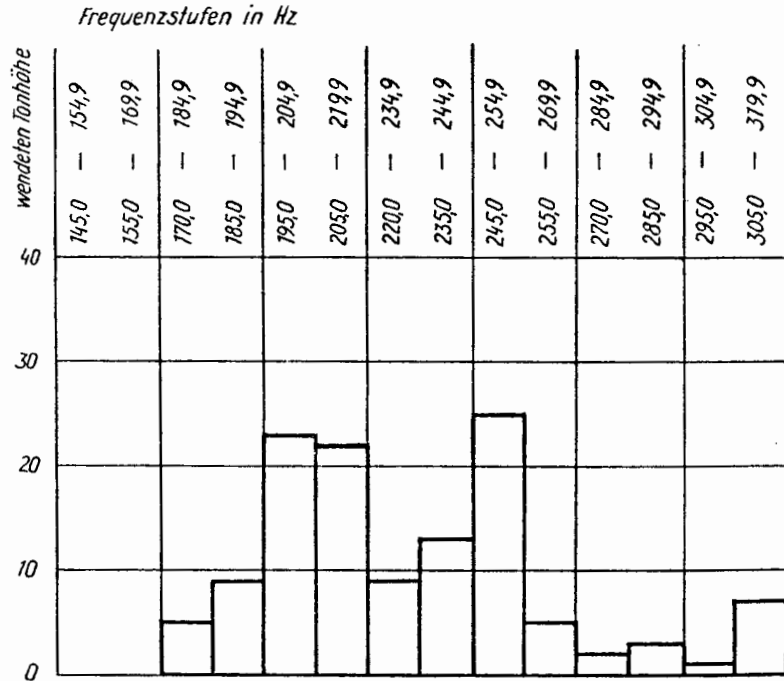


Abb. 3. Darstellung der Häufigkeit der verwendeten Sprechhöhen — Frauen Satz (Ansage des Datums)

schnittswert der registrierten Frequenzen sowie die am häufigsten verwendete Sprechtonhöhe größere Unterschiede aufweisen. Auf dem folgenden Säulendiagramm wird ein derartiges Beispiel angeführt.

Die Abb. 3 zeigt die Häufigkeit der Tonhöhenverwendung gleichfalls bei der Ansage des Datums. Diese weibliche Vp sprach unrhythmisch in relativ großen Tonschritten. Die Sprechmelodie hatte einen sogenannten „singenden“ Charakter. Dieser wurde durch die dialektgebundene Melodisierungsweise hervorgerufen. Die Säulen zeigen eine unsymmetrische Anordnung, das Erfassen des HSTs war recht schwierig. Die Abb. 4 zeigt die Bestimmung des individuellen physiologischen HST-Bereichs (mit Angabe der Durchschnittswerte für Männer in Noten). Er befindet sich im unteren Drittel des Singstimmumfangs und dieser wiederum im Rahmen des potentiellen Stimmumfangs. Der individuelle physiologische HST-Bereich, der von der kritischen HST-Höhe und der kritischen HST-Tiefe begrenzt wird, enthält die Spannungshöhe, die Lösungstiefe und den Normalsprechton. Dieses Untersuchungsergebnis konnte durch die Apperzeption der individuellen HST-Höhen bei unterschiedlichen Sprechleistungen der Versuchspersonen ermittelt werden.

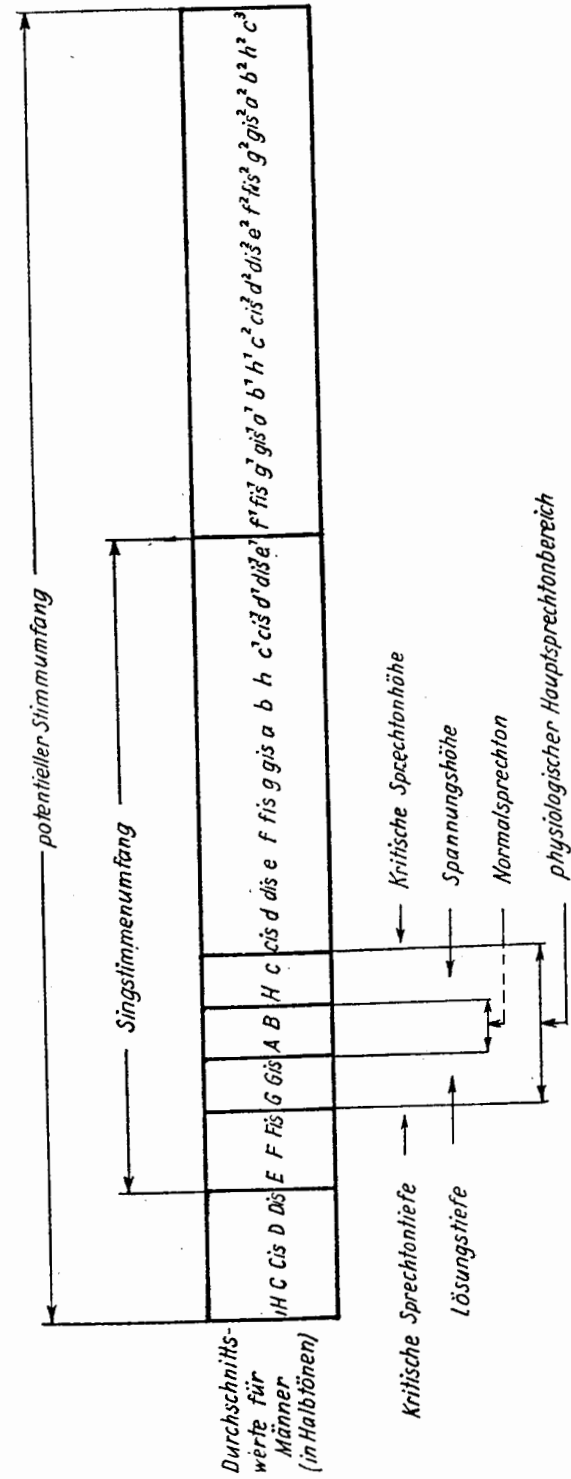


Abb. 4. Darstellung der Aufteilung der individuellen Stimmbereiche

L'UTILISATION DE LA CHRONOLOGIE RELATIVE DANS LA DESCRIPTION STRUCTURALE DE L'ÉVOLUTION PHONOLOGIQUE DES LANGUES

VLADIMÍR HOŘEJŠÍ

Dans ma brève communication, je voudrais soumettre à votre examen critique quelques-unes de mes idées sur l'utilisation de la chronologie relative dans la description structurale de l'évolution phonologique des langues, idées dont j'ai exposé l'essence dans le premier volume des Travaux linguistiques de Prague (1964) et dans le volume publié en l'hommage du linguiste roumain Alexandru Rosetti (Bucarest, 1965).

La description de l'évolution phonologique d'une langue (de même que la description de son évolution morphologique, syntaxique, etc.) n'est possible, à notre avis, qu'après l'achèvement de deux opérations préalables:

1° l'établissement de la chronologie relative des changements linguistiques isolés;

2° l'évaluation de ces changements au point de vue du système linguistique respectif.

La première des deux opérations a été déjà assez recherchée de son côté méthodique et assez vérifiée par son application aux matériaux linguistiques concrets pour qu'il me semble utile d'en parler davantage. Qu'il me soit donc permis de procéder sur-le-champ à l'investigation — du côté méthodique et en vue des résultats à obtenir — de l'autre opération. Je ne m'occuperai, naturellement, que des changements affectant le système phonologique de la langue (tout en laissant entendre que des procédés analogues pourraient se faire valoir lors de l'investigation des autres plans linguistiques) et je ne parlerai pas non plus de l'influence que peuvent avoir les changements survenus à l'intérieur d'un plan linguistique sur les faits d'un autre plan.

Pour pouvoir donc évaluer les changements linguistiques au point de vue de l'influence qu'ils peuvent exercer sur le système phonologique de la langue (ce procédé présuppose leur rassemblement antérieur et l'établissement de leur ordre chronologique), il faut les caractériser proprement en déterminant, dans chaque cas à part, le type de changement accompli. Il va de soi qu'il faut prendre en considération, au moins préalablement, tous les changements phoniques subis par la langue sans se soucier d'avance s'ils sont d'ordre phonologique ou purement phonétique. Dans le plan phonique de la langue, les changements peuvent être de trois types:

1° changement d'un élément sonore en un autre;

2° disparition d'un élément sonore (ce qui équivaut à son changement en un zéro phonique);

3° apparition d'un élément sonore (ce qui équivaut au changement du zéro phonique en un élément sonore).

Ce que j'appelle „élément sonore“, peut être — tant en fonction de ce qui change qu'en fonction de ce qui est le résultat du changement — un son unique ou bien un groupe de sons. Chacun des trois types de changement énumérés peut être inconditionné, si le changement respectif se fait mécaniquement dans tous les lieux de l'apparition du son qui change, ou conditionné, si le changement se réalise ou non en fonction de conditions externes, par ex. en fonction de l'entourage phonétique de l'élément sonore respectif. Ainsi, le nombre des types de changement s'élève théoriquement à six, mais pratiquement à cinq par l'absurdité de l'apparition inconditionnée d'un élément sonore à partir du zéro phonique.

Après avoir été caractérisé de la manière indiquée et classé sous un des cinq types mentionnés, tout changement phonique devrait être examiné au point de vue de l'influence qu'il a exercée sur le système phonologique de la langue ou, si vous voulez, au point de vue de sa répercussion dans le corps de ce système.

Une telle investigation — y compris les résultats qu'elle donne — est loin d'être complète si elle ne va pas — comme il est d'usage dans les publications existantes — au-delà des changements qui surviennent ou non dans l'inventaire des phonèmes et dans le système des oppositions phonologiques. Cependant, la réalité est bien plus complexe. Des changements peuvent survenir non seulement dans l'inventaire des phonèmes et dans le système des oppositions phonologiques, mais aussi dans la réalisation phonétique des phonèmes (y compris leur réalisation par différentes variantes et les rapports entre celles-ci), dans le contenu phonologique des phonèmes, dans leur distribution à l'intérieur du mot, dans leurs facultés combinatoires et, enfin, dans leurs fonctions démarcatives.

Une telle extension du domaine à rechercher, si considérable qu'elle soit, ne saurait toutefois être unique, mais demande à être complétée par une autre. C'est que les changements survenus dans le plan phonique de la langue affectent, sous tous les rapports mentionnés ci-dessus, non seulement les sons et les phonèmes qui subissent directement les changements, mais aussi les sons et les phonèmes environnants; ainsi, par ex., la disparition d'un son ou phonème au début ou à la fin du mot peut changer considérablement les qualités du son ou phonème voisin, qui, maintenant, devient lui-même initial ou final respectivement.

Il résulte de toutes ces considérations qu'une description exhaustive de l'évolution du système phonologique d'une langue (et, à plus forte raison, de l'évolution d'un système linguistique entier) est extrêmement compliquée. Mais les problèmes scientifiques qui paraissent à première vue insurmontables sont progressivement maîtrisés, surtout dans les dernières années. Il est important de les connaître et d'approcher pas à pas de leur solution.

DISCUSSION

Haudricourt:

La différence entre la conception classique de la chronologie relative et la conception structurale réside dans l'utilisation de la neutralisation. La conception classique classe des changements phonétiques ponctuels (dans le temps).

Dans la conception structuraliste, une neutralisation peut durer longtemps et être simultanée de plusieurs changements successifs (exemple: l'apparition des voyelles Umlaut en germanique, commençant avec la disparition des YOD, et se terminant avec la disparition des I finals).

THE DISAPPEARING POST-VOCALIC [r] OF „GENERAL AMERICAN“ ENGLISH SPEECH

ALAN W. HUCKLEBERRY

The International Phonetic Alphabet consonant-vowel system has been accepted throughout the United States. It has been usable as a generalization and also as an initiation of the undergraduate university student to the study of the sounds of languages. The IPA has not always been applicable, both in description and insight, to American speech, and especially to that speech known as “General American.” One concept in particular is not applicable, and that is the description of an [r] sound. It shall be the purpose of this paper to present a point of view in regard to one aspect of the [r] phoneme. This one aspect is the delineation of the post-vowel [r], referred to in this paper as the post-vocalic [r] sound. It shall present the probability that no such sound exists and that other sounds are in its place.

The philosophic base which started this semi-scientific study is the “observational rule” of all languages that *two (or more) consonants adjacent to each other in the same syllable must be either both voiced or both voiceless*. This particular “rule” is probably based on the physiological fact that such consonants become blended into *one* sound and therefore cannot have the qualities of one being voiced and the other voiceless. Examples to illustrate the “rule” are such as “stove” [stov], “reached” [ritʃt], “space” [spers], “kites” [kitts], and “beds” [bedz]. Nevertheless, the rule does have some rather obvious exceptions at first view. The *first* is that a voiced nasal can be combined with a voiceless consonant as in [smɒl], or [went], and [tærk]. This exception needs further research. A *second* exception is that the voiced [l] sound *can* be combined with voiceless consonants. My research, as well as others, tends to show that the [l] is not *always* voiced, as described in the IPA, and that the pre-vocalic [l] in combination with a voiceless consonant, as in “play” [pleɪ] or “sleep” [slɪp], is also voiceless. Further research is needed, however, on the post-vocalic [l] as in “help” [hɛlp]. A *third* exception is the pre- and post-vocalic [r]. My research, as well as that of others, leads to rather definite evidence that the voiced pre-vocalic [r] sound, when in combination with a voiceless consonant, is a voiceless [r]. This can be shown with either the [r] sign or by using the [ɹ] symbol. I prefer the latter. Examples of the pre-vocalic voiceless [ɹ] concept are “train” [tuɛɪn], “crowd” [kɹɑud], “thread” [θɹɛd], and “pray” [pɹeɪ]. Laboratory research at Ball State University illustrated, what others have inferred, that adjacent consonants *do* blend together and become *one* consonant. Thus, the two sounds of [k] and [r] in “cream” become one voice-

less sound, i.e. "cream" [kɹim]. Tapes of rapid speech of such a speech-word as [kɹim] can be cut to eliminate the [m] and the [i]; but if the [k] is removed, the [ɹ] disappears with it. Even the three sounds preceding and following a vowel, as in the word "scratched," [skɹætʃt] follow the same principle. Thus, the observational rule usually holds for adjacent consonants being either all voiced or all voiceless. The only exceptions seem to be certain [l] and certain nasal combinations.

It is the *post-vocalic* General American English [r] that is now the subject of analysis for the remainder of this paper. What is the nature of the [r] that follows a vowel but is adjacent to another consonant? Examples are such as "card" [kɑrd], "cart" [kɑrt], "chord" [kɔrd], and "court" [kɔrt]. If the observation "rule", previously mentioned, holds true, then a voiced [r] would need to be followed by another voiced consonant, such as [d], and not by a voiceless consonant, such as [t] in the previous examples. The inconsistency is intriguing—and so the study—and so this paper.

First, a brief "aside" in regard to the division of syllables in General American English in regard to the [r] sound. Where the Britisher, the Eastern American, and the Southern American has a choice, he seems to divide his syllables in order to permit the [l] and [r] sounds to become *pre-vocalic*. The General American speaker, on the other hand, is likely to place the [l] and [r] in a *post-vocalic* position or even change the complete word pattern. Examples of the Southern, Eastern, and British speakers' style would be "silly" [sɪ-lɪ], "Alice" [æ-lɪs], "salad" [sæ-lɪd], "parade" [pə-ri-d], "carrot" [kæ-rət], and "perish" [pɛ-rɪʃ]. The General American speaker is likely to say [sɪl-ɪ] not [sɪ-lɪ], [æ-l-əs] not [æ-lɪs], [sæl-əd] not [sæ-lɪd], [pə-ri-d] not [pə-ri-d], [kæ-rət] not [kæ-rət], and [pɛ-rɪʃ] not [pɛ-rɪʃ]. This rather consistent shift gives the General American speaker many additional *post-vocalic* [r] sounds. *But then*, considering our observational rule, how can that "voiced" [r] be blended with a "voiceless" consonant as in "cart" [kɑrt]? Is the [r] the same as the [ɹ] in the *pre-vocalic* position? Or does the [r] exist at all in the *post-vocalic* position? The evidence now points to the idea that the voiced *post-vocalic* [r] does *not* exist and is not blended with a voiceless consonant, i.e. [kɑrt]. Research for years has hinted that it takes on a "vowel quality." What does happen?

Laboratory study appears to support the original hypothesis of this paper that the *post-vocalic* [r] *does not exist*. Rather, two phonetic phenomena appear to take place:

1. If *what we have thought* were the *post-vocalic* [r] appears in a syllable which has the *primary* accent or appears at the *end* of a phonetic expression before a pause, it becomes a vowel [ɹ̥]. For example, if a General American speaker says, "He wrecked his car", he will say it as [kæɹ̥], not as [kɑr].

2. If the *post-vocalic* [r] appears in a syllable which has the *secondary* accent, it then becomes a part of the previous vowel. It actually produces—or is—a *new* vowel. Thus, if a General American speaker says, "His car was wrecked," he will

say it as [kæɹ̥]. The symbol following the [k] is fabricated and is certainly *not* included in the IPA!

The following will further illustrate the results of this study:

Word	Primary Accent or Before a Pause	Secondary Accent
1. clear	[klɹ̥ə]	[klɹ̥]
2. mayor	[meɹ̥]	[meɹ̥]
3. cares	[keɹ̥z]	[keɹ̥z]
4. cart	[kæɹ̥t]	[kæɹ̥t]
5. stores	[stɔɹ̥z]	[stɔɹ̥z]
6. sure	[ʃʊɹ̥]	[ʃʊɹ̥]
7. tower	[taʊɹ̥]	[taʊɹ̥]
8. tire	[taɪɹ̥]	[taɪɹ̥]

The material just stated was "tested out" in three ways. It was *first* tested by a process of clipping magnetic tape recordings of rapid speech. Under primary accent, the tape could be clipped off so that the [ɹ̥] was removed and the previous vowel remained. Under secondary accent, the [r] concept could *not* be cut off the tape without removing the preceding vowel with it. *Second*, a frequency analysis was made and found that only vowel frequencies—and not the *pre-vocalic* [r] frequencies—were present when the so-called *post-vocalic* [r] was present in either primary or secondary accented positions. *Third*, sonograms were made on a sound spectrograph and studied in relation to other consonants and vowels. The *postvocalic* [r] came out consistently as a vowel or vowel-blend when studied in terms of formants. There appears to be enough evidence to call for further investigation. The evidence presents valid doubts regarding the conventional IPA [r] concept.

What are some implications if such new studies do appear to be valid? Some future questions could be:

1. Do we have a new set of vowels in [ɑ̥], [ɹ̥], [ʊ̥], etc.?
2. Is there a need to have some new vowel diagram that is a variation of the IPA? The Americans, in general, use the quadrangle of the IPA and I understand that some of the French use the triangle. I am developing one that is curved to match the curvature of the hard and soft palate.
3. Would this lend to a new approach to speech correction for those persons who cannot make the *post-vocalic* [r]?
4. Is this a cementing of the theory of language communication that the complete syllable is the acoustic signal for conveying meaning of thought?
5. Is there a need for having an international congress to propose ideas for revision, modernization, and standardization of our IPA? If so, I would be willing to work for such a conference.

DISCUSSION

McDavid:

1. The term "General American" has little validity. In Kurath's *Word Geography of the Eastern United States* (1949) it is pointed out that the major dialect regions of the Eastern States are a) Northern—New England, the Hudson Valley, and derivative settlements westward; b) Midland—Pennsylvania and derivative settlements to the South and West, including large areas of all of the Southern States; c) Southern—the older plantation settlements from Maryland south. A. H. Marckwardt, A. L. Davis, and Virginia McDavid have shown that in the Great Lakes basin there is a major dialect boundary crossing Ohio, Indiana and Illinois; Harold B. Allen has found a similarly sharp boundary in the Upper Midwest west of the Mississippi. These findings have been replicated by such independent investigators as the late C. K. Thomas.

2. The boundary between "consonantal" and "vocalic" (-r-) in such words as *carrot* was shown by Thomas, some years ago, to cross Ohio, Indiana and Illinois.

3. As a footnote: in many southern and south Midland speakers, the strong consonantal intervocalic (-r-) is lost in *carrot* etc. often by speakers who have a strongly constricted (-r) in *cart* and the like.

Hill:

I think that Prof. Huckleberry might consider whether his rule that there can be no succession voiced-voiceless (or the reverse) consonants in the same syllable should not be revised to say that the rule does not apply to consonants where voice quality is not a distinctive feature, as with liquids and nasals.

Vachek:

Prof. Huckleberry may be perfectly right in speaking about the disappearance of /r/ from the phonetic point of view, but from the phonological point of view in the described type of American English pronunciation /r/ is preserved as a phoneme. In the New England (and the Southern) type of American English pronunciation, however, the /r/-phoneme really disappears, just as in the Southern British pronunciation.

LA FRÉQUENCE DES FINALES VOCALIQUES DANS LES PARLERS RHÉTOROMANS

MARIA ILIESCU

En considérant le caractère vocalique ou consonantique des finales des mots comme un trait typologique non dépourvu d'importance, je voudrais présenter ici la situation des finales, spécialement la situation des finales vocaliques dans les parlers rhétoromans, telle qu'elle se dégage d'une première recherche faite à l'aide de la méthode statistique. En ce qui est du terme „rhétoromans“ il me semble nécessaire de corriger en ajoutant „des parlers dits rhétoromans“ car je crois être en accord avec la plupart des spécialistes en affirmant que le rhétoroman n'est qu'une notion conventionnelle pour un grand nombre de dialectes et patois parlés, on le sait bien, dans une partie de la Suisse et de l'Italie du Nord, dialectes dont la comparaison structurale et typologique n'a pas encore été faite. J'ai choisi pour ma recherche les dialectes les plus représentatifs: le sursilvain et le haut engadinois pour le groupe de l'ouest, le dialecte du Val Gardena pour le ladin dolomitique et le frioulan dans la variante de Udine.

La statistique a été faite sur le même texte, du point de vue du contenu, pour tous les quatre dialectes, à savoir: la traduction dans les idiomes respectifs des chapitres XX jusqu'à XXIV des Actes des Apôtres qui ont été à ma disposition grâce à la série: G. de Poerck et L. Mourin, *Introduction à la morphologie comparée des langues romanes* éditée par la Faculté des Lettres de Gand. J'ai consulté en manuscrit le texte frioulan. Comme une grande partie des mots terminés en voyelles sont des prépositions et des noms féminins provenant de la première déclinaison latine, j'ai cru absolument nécessaire de choisir un texte avec le même contenu, bien que pour les premiers deux dialectes il s'agisse de textes anciens tandis que pour les deux derniers de traductions contemporaines. Une statistique comparative faite pour les deux textes anciens avec des textes contemporains des mêmes dialectes m'a donné l'assurance que du point de vue des finales il n'y a pas de différence considérable entre l'époque ancienne et contemporaine.

D'un autre côté une comparaison faite entre les données obtenues pour le premier et le second millier d'occurrences a démontré que le résultat statistique ne présentait que des différences négligeables ainsi que j'ai cru suffisant d'étendre l'analyse sur 5000 mots pour chaque texte. Il me faut encore préciser que j'ai considéré comme une seule occurrence les mots, surtout articles, prépositions et conjonctions dont les

signes orthographiques (apostrophes, traits d'union, etc.) marquaient l'unité du mot respectif avec son antécédent.

Les chiffres obtenus, en pourcentage, sont ceux écrits sur le tableau

Dialecte	Finales vocaliques	Finales consonantiques
Sursilvain	38,70	61,30
Engadinois	40,10	59,90
Frioulan	57,50	42,50
Val Gardena	63,50	36,50

Les différences, on le voit, sont parfois substantielles. Tandis que le sursilvain et l'engadinois se caractérisent par des finales consonantiques, le dialecte de Val Gardena présente un grand nombre de finales vocaliques. (Il est intéressant à remarquer que le pourcentage des finales vocaliques propre à ce dialecte est presque le même que celui trouvé par mon collègue, C. Maneca, pour le roumain: 65,47 %.) Le frioulan occupe une position intermédiaire.

La tendance vers syllabe finale ouverte ne croît pourtant pas proportionnellement de l'ouest à l'est comme il paraît à première vue, puisque le frioulan, le plus estique des idiomes analysés, est moins vocalique quant aux finales que le dialecte du Val Gardena. Il me faut toutefois attirer l'attention que dans cette recherche j'ai fait abstraction des voyelles finales nasales, étant donné que j'ai été obligée par les circonstances d'analyser seulement des textes écrits.

Du point de vue qui nous intéresse, les dialectes de l'ouest (sursilvain et engadinois) font en tout cas nettement groupe à part.

La fréquence de chaque voyelle varie aussi considérablement dans chaque dialecte. Il est à noter que j'ai fait abstraction des diphtongues et triphongues et que je n'ai pas séparé les finales accentuées (peu nombreuses) des finales non-accentuées.

Dialecte	a	e	i	o	u	ü
Sursilvain	75,33 %	2,66 %	12,33 %	4,33 %	5,33 %	
Engadinois	53 %	17 %	12 %	13,49 %	5,41 %	1,10 %
Val Gardena	28 %	31,33 %	20,75 %	11,10 %	8,31 %	
Frioulan	14 %	31,27 %	30,53 %	20,21 %	3,99 %	

On observe tout d'abord la grande prépondérance du *a* et du *e* vis-à-vis du *o* et *u* et le décroissement du *a* en faveur du *e* de l'ouest à l'est. Il faut naturellement tenir compte de la valeur phonétique de chaque lettre car la différence entre l'orthographe et le phonétisme est parfois substantielle. Ainsi y a-t-il une grande identité de prononciation entre le *a* et le *e* atone en sursilvain et en engadinois. Ces lettres représentent dans les deux dialectes un son [ə] intermédiaire entre *e* et *a* avec beaucoup

de variantes, son habituellement moins fermé quand il est écrit *a* et plus fermé quand il est écrit *e*. Pour le frioulan il faut remarquer que le grand pourcentage du *e* vis-à-vis du *a* est spécifique pour le dialecte de Udine car dans d'autres patois les finales des noms féminins provenant de la première déclinaison latine présentent *a*. Mais dans ce dialecte il s'agit aussi généralement d'un son [ə] avec beaucoup de variantes. Le trait le plus caractéristique du frioulan, du point de vue des finales vocaliques, est la fréquence du son *i*. Toujours pour ce dialecte il faut ajouter que le pourcentage relativement élevé du *o* dans le texte analysé par nous provient des italianismes assez fréquents dans le manuscrit que j'ai été contrainte d'utiliser.

En conclusion l'analyse statistique entreprise nous permet de dégager un trait qui sépare les dialectes dits rhétoromans (à savoir le caractère consonantique des finales du groupe de l'ouest et le caractère vocalique des finales du groupe central et de l'est) et un autre trait qui les unit (la prépondérance des voyelles finales neutres et antérieures).

J'espère avoir démontré implicitement l'utilité de la méthode statistique appliquée à la phonétique et son apport pour l'élucidation de certains problèmes litigieux de la dialectologie pour la typologie des différents idiomes.

DISCUSSION

Wittoch:

Le travail de Mme Iliescu tend vers la précision qu'elle a atteinte déjà en grande partie. Il y a, cependant, une certaine différence, par ex., entre le haut et le bas engadinois, à ce point de vue-là (le „a“ final inaccentué du haut engadinois pourrait être comparé avec le „ă“ du roumain!), pour ne pas parler des „dialectes“ rhétoromans du Haut Rhin. Les idiomes rhétoromans attendent encore des recherches plus détaillées qu'on n'ait pu faire jusqu'ici.

PERCEPTION OF STRESS BY CZECH LISTENERS

P. JANOTA*

The purpose of the experiments to be described was to investigate the relation between synthetic speechlike stimuli varying in their physical dimensions and the auditory perception of these stimuli as stressed or unstressed patterns. The evaluation of the examined phenomena is based on listening tests, presented to groups of listeners of Czech nationality.

Irrespective of the function of stress in the Czech language, it may be assumed that Czech listeners are able to discriminate between stressed and unstressed parts of an utterance, one syllable of which is stressed and the other one unstressed. Generally, the results of tests with simplified synthetic speech cannot yield a direct explanation of the relations in natural speech. Nevertheless, they can fairly sensitively indicate the trends existing in natural speech and consequently give rise to reasonable hypotheses.

For the listening test a synthetic combination of two vowels and two consonants was chosen, the order of the sequence being CVCV. This type of sound combination is very frequent in Czech; as a disyllabic word it can occur only with the stress on the first syllable, as a disyllabic rhythmical structure it can have the stress on the first or on the second syllable. The listeners had to decide which of the two synthetic syllables they heard as stressed. It was thus desirable that sounds inserted for the "C" and "V" elements did not form any common Czech word. Accordingly, an "s" was inserted for both C-elements and an "e" for both vowels. The combination "sese" was then used in all items of the listening test, so as to keep the acoustic structure of stimuli — with the exception of variables — constant throughout the whole test. It is supposed that in this connection two further assumptions may be made, *viz.* (1) that no considerably different results could be expected, if other speech sounds were inserted for the V and C and (2) that a different number of syllables in the test items would not yield results fundamentally different from those for disyllabic stimuli. In the present test, three variables were introduced: (1) the sound pressure of the second vowel, resulting in changes in perceived loudness (*I* intensity), (2) the fundamental frequency of the second vowel, giving rise to changes of pitch (*F* —

* Institute of Phonetics, Charles University, Prague.

frequency) and (3) the duration of the second vowel, perceived as a varying quantity of the vowel (*D* — duration); in several items, only the duration of the second consonant was changed. All variable dimensions were changed on five levels (± 3 and 9 dB, \pm one and three semitones and five logarithmic steps in the range 80—200 ms for the I, F, D dimensions, respectively).

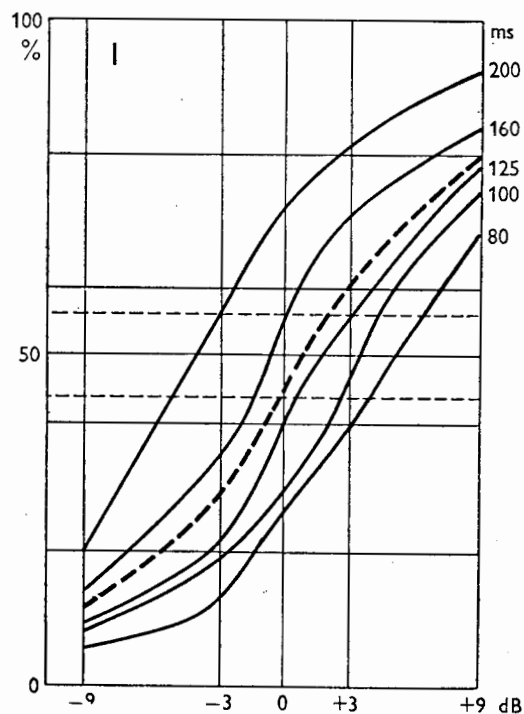


Fig. 1.

The artificial sounds were produced by means of a steady-state synthesizer, consisting of four parallel-connected formant circuits with controllable central frequencies, bandwidths and amplitude levels. The test items (130 in all) were then randomized, stored on tape and presented to 174 subjects who were all young native speakers of Czech. In test evaluations, over 22 000 judgments were subjected to computations. The results were then transferred into tables and graphs, only four of which will be shown here. In the first graph the change of judgments "stressed" is plotted as density function of intensity for respective levels of duration, the remaining dimension (frequency) being kept constant (Fig. 1). In addition, a curve for all duration values averaged is shown. Above and below the 50 % line an "interval of uncertainty" is drawn, computed as the critical difference between chance identifications and identifications required for significance at the 5 % level of confidence. Similarly, the second graph (Fig. 2.) was plotted for changes in duration (on five intensity levels) and in the third graph (Fig. 3.) percentage frequencies of judgments

vs. frequency in cps for five intensity levels is shown. It is apparent that the changes in the frequency dimension influence the judgments of stress with much less weight and consistency than the other two factors involved. The distribution of the points seems to show fairly convincingly that there is a sudden change in judgments between items with both test syllables of equal fundamental frequency and those where the

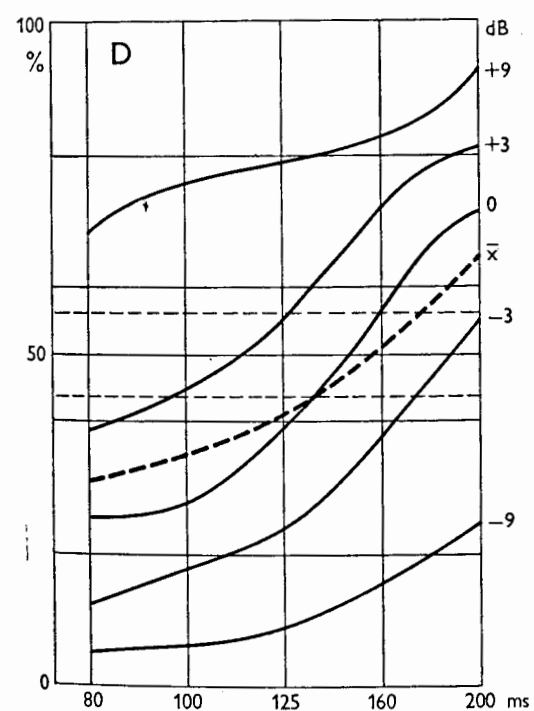


Fig. 2.

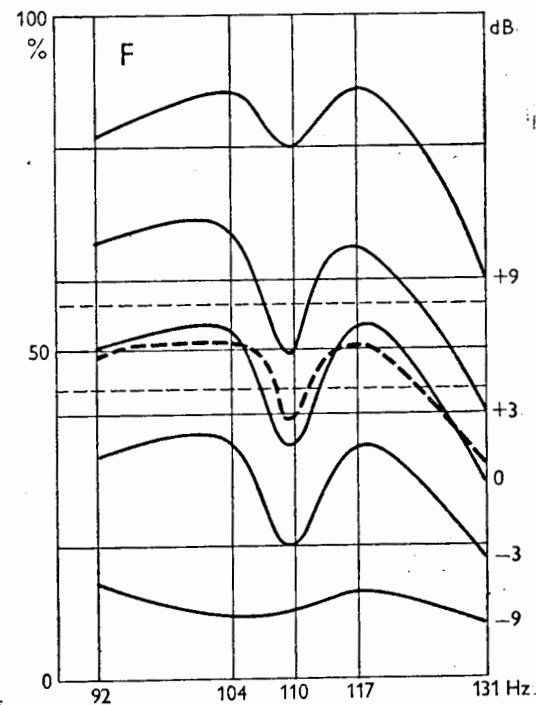


Fig. 3.

fundamental frequency was changed slightly. Slight changes (one semitone) are associated with an increase of judgments in the direction "stressed". By introducing somewhat greater changes of fundamental frequency, a reversed effect is evoked, and the number of test syllables rated as stressed diminishes in both cases.

In the last graph shown here (Fig. 4.) data of judgments changing simultaneously and in the same direction in two and three physical dimensions are given. The total mean value of all judgments is included in the form of a horizontal line. Its position below the 50 % line shows that in borderline cases Czech listeners more often perceive the stress on the first syllable. In combined plots of the graph shown, the expected effect of a simultaneous change in two or three dimensions is obvious: the central portion of both curves is very steep and even the minor effect of frequency changes adds to the abrupt swing between the "stressed" and "unstressed" areas.

According to traditional views on "dynamic accent" in Czech, the intensity factor was often presented as the fundamental one and the influence of durational changes

denied. Experiments in the perception of stress in various languages, some of them using also synthetic material, show, in a majority of cases, the outstanding perceptual importance of fundamental frequency changes. This was shown e.g. by Mol and Uhlenbeck, Fry, Strevens, Bolinger, Lieberman, Malmberg, Rigault, Jassem, Isačenko and Schädlich and others. In the present experiment no such prevailing impor-

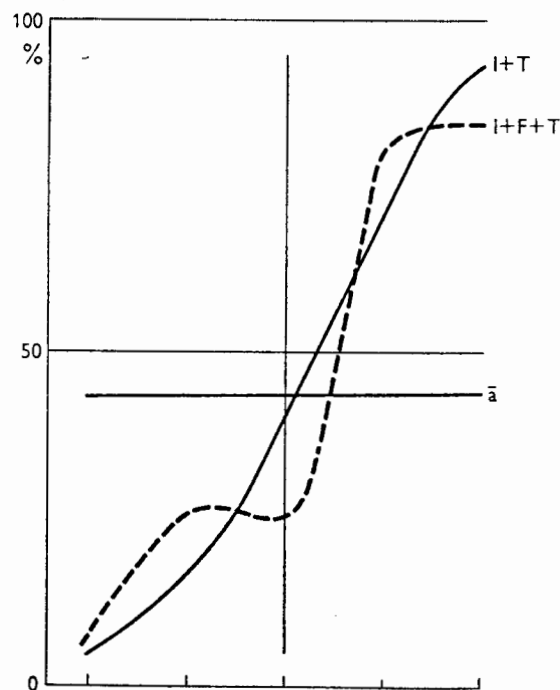


Fig. 4.

tance of fundamental frequency changes was experienced by Czech listeners. (By detailed statistical analysis it has been demonstrated that this trend was apparent in a group of Czech students of French.) A limited investigation into the relative effect of durational changes of the fricative sound used as a C-element in test items shows that a decrease in duration of the consonant causes a slight increase in the number of "stressed" judgments and an increase in duration a more pronounced one. In both cases the effect is weaker than the influence of changes in the vocalic element. The effect of sequence in pairs of adjoining stimuli is statistically significant, i.e. an influence of the immediately preceding item on the following one was stated, working in the direction of contrast. A possible equivalent of this effect in natural speech signal cannot be excluded from consideration. It is hoped that the results reported may contribute to a better knowledge of the auditory evaluation of stress in Czech.

DISCUSSION

Fischer-Jørgensen:

a) As Czech has phonologically relevant vowel length also in unstressed syllables, one should expect duration to be less important for the perception of stress than in other languages. Has such a comparison been possible? b) Does the amount of increase in frequency, duration and intensity required for perceiving one syllable as more stressed than be neighboring syllable differ for first and second syllable?

Rischel:

I wish to ask whether the segments constituting the test stimuli were entirely steady-state sounds switched on and off abruptly. One might consider the possible relevance of the type of onset and offglide for the judgment of stress placement.

Sivertsen:

Were your listeners linguistically trained or naive? What instructions were given to them they were asked to identify the stressed syllable? What did your listeners mean by 'stress'?

Janota:

Ad Fischer-Jørgensen: Principally, a comparison of listening tests presented to native speakers of various languages would be possible and quite useful. In that case, however, the test material would have to fulfil comparable conditions of acceptability for both or all languages investigated. As yet, a similar effect in the frequency domain could be demonstrated indirectly on results of the test presented to Czech students of French. In this group, a greater relative importance of fundamental frequency changes for the perception of stress was apparent.

In our test, the amount of increase in frequency, duration and intensity required for perceiving one syllable as more stressed than the neighbouring one does not differ for the first and second syllable of the test items used, except for the fact that the total mean value of all measurements shows a slight shift in the direction to judgments of 'stress on the first syllable'. (Cf. Fig. 4.)

Ad Rischel: The segments constituting the test items were not switched on and off abruptly; besides, all the stimuli were stored on tape. Two identical test tapes were used for all the tests.

Ad Sivertsen: The groups of listeners consisted mainly of first year students of the philosophical faculties in Prague and in Brno, so that they were not completely linguistically naive, though they were not specially trained for the test. The instructions to the test were recorded on the test tape so as to exclude a possible influence of a non-uniform formulation of the task on the results of the experiment; the listeners were asked to underline the stressed syllables of all items. No explicit explanation was given as to what was meant by stress (rather, we expected the answer from the listeners), though the listeners' linguistic knowledge could not, of course, be eliminated. It was expressly mentioned in the test instructions that the test items were synthetic, but the purpose and composition of the test was not explained.

INFORMATION BY TONES AND SINGLE VOWELS

L. KAISER

As the results of these researches were published in relatively unknown periodicals, partly even in Dutch, it may be of some interest to summarise these here in a circle of professional phoneticians, the more so as they seem to fit into the general programme.

The material consisted of two collections of combinations of three tones, in which duration and intensity had been kept constant as much as possible. The first collection consisted of 49 sequences, giving rise to associations and afterwards leading to a judgment concerning the emotional and linguistic information.

The second collection implied 2535 sets of three tones, all lying between *c*, *c'* and *c''*. Here only a judgment was demanded concerning the acceptability of the tune for three geographic names, as a statement, a question or a call. If in this case emotion was judged to be a condition for a certain possibility, this was simply noted.

Besides we had at our disposal the records of three vowels, *a*, *o*, *e*, performed by 3 male and 7 female students, judged by 50 students as to the affect they meant to transfer: cheerfulness, disgust, kindness, enthusiasm, grimness and sadness.

It was striking that concerning the emotional information of the 49 complexes of three tones a high degree of concordance existed, whereas concerning the linguistic information the opinions differed.

It appeared that the first interval had a preponderant rôle in the expression of emotion.

The linguistic category: statement, question, call, seemed to be determined especially by the second interval.

As to the judgments concerning the three geographical names fitting to a tune of three tones, the results were regular to a high degree. Generally spoken the positive results were found in one single quadrant. For the first word (Apeldoorn) this was the one of two negative intervals, for the second word (Ootmarsum) a first positive interval was combined with a second negative one, while for the third word (Oostmahorn) a field was found covering two quadrants partly, in which a rising second interval was combined with a first one that was either falling or rising. This ambiguity found its cause in a relatively strong secondary accent on the first syllable.

A special behaviour showed the territories in which one of the intervals was zero, this favouring the impression of a call.

Mr Van Gelder tried to formulate the importance of the range of the intervals and that of absolute tone value.

The second interval presented 169 possibilities, eventually leading to the interpretation of one, two or three names, as a statement, a question or a call.

Tracing a diagonal, we find at the right the falling and at the left the rising intervals. In this latter half statements are almost absent, whereas the right one shows a mixture of questions and statements. C or cis as an ultimate tone frequently results in a statement, but it is striking that especially here also many questions occur.

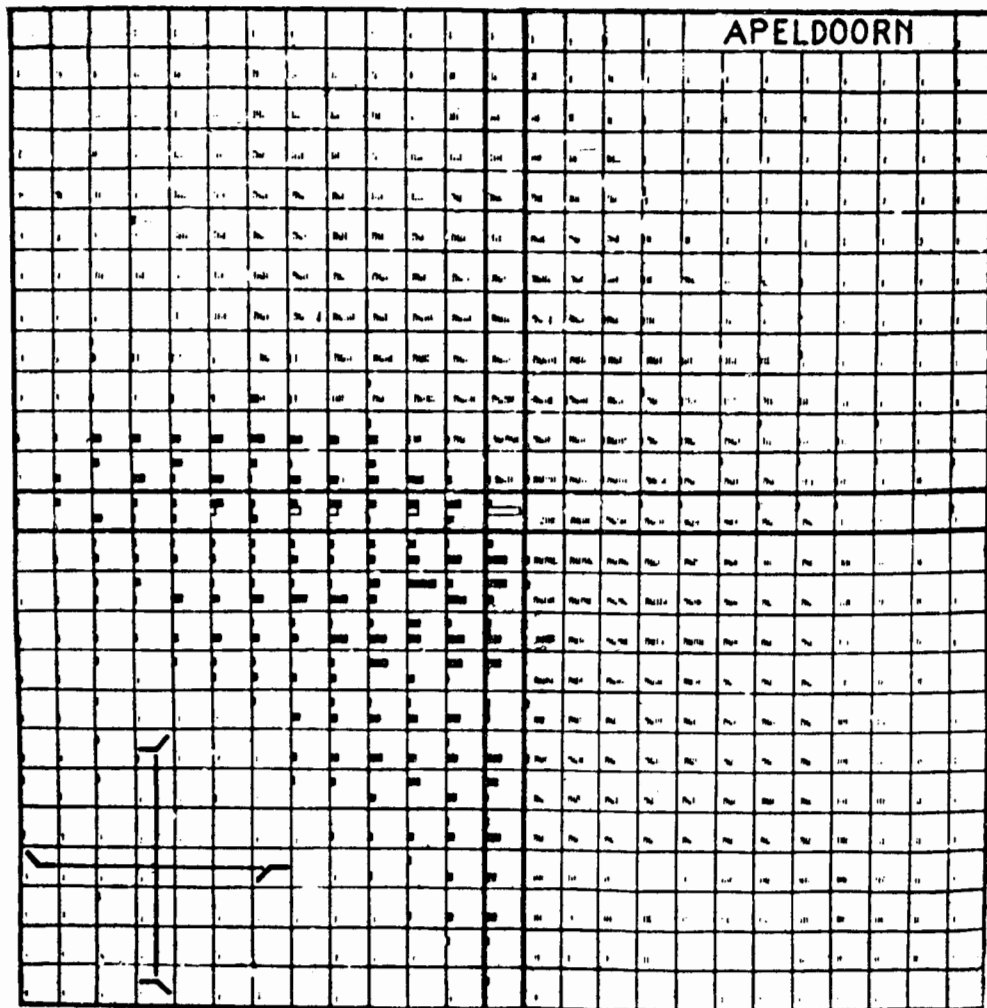


Fig. 1.

Questions show maximal values in rising intervals encompassing two halftones. Falling intervals encompassing about six halftones work also as questions.

A call was especially recognized if the ultimate tone was lying between c¹ and g.

Fis as the first tone of the first interval gave a remarkably high number of statements for all three words, demonstrating an exceptional linguistic role of the first interval.

Of 127 tunes on c giving a signal, this concerned 70 times a statement, 53 times a question and 4 times a call.

In the research with the vowels the human vocal sounds gave rise to an extremely prompt mimical reaction, followed by free associations which often were highly emotional and accompanied by gestures.

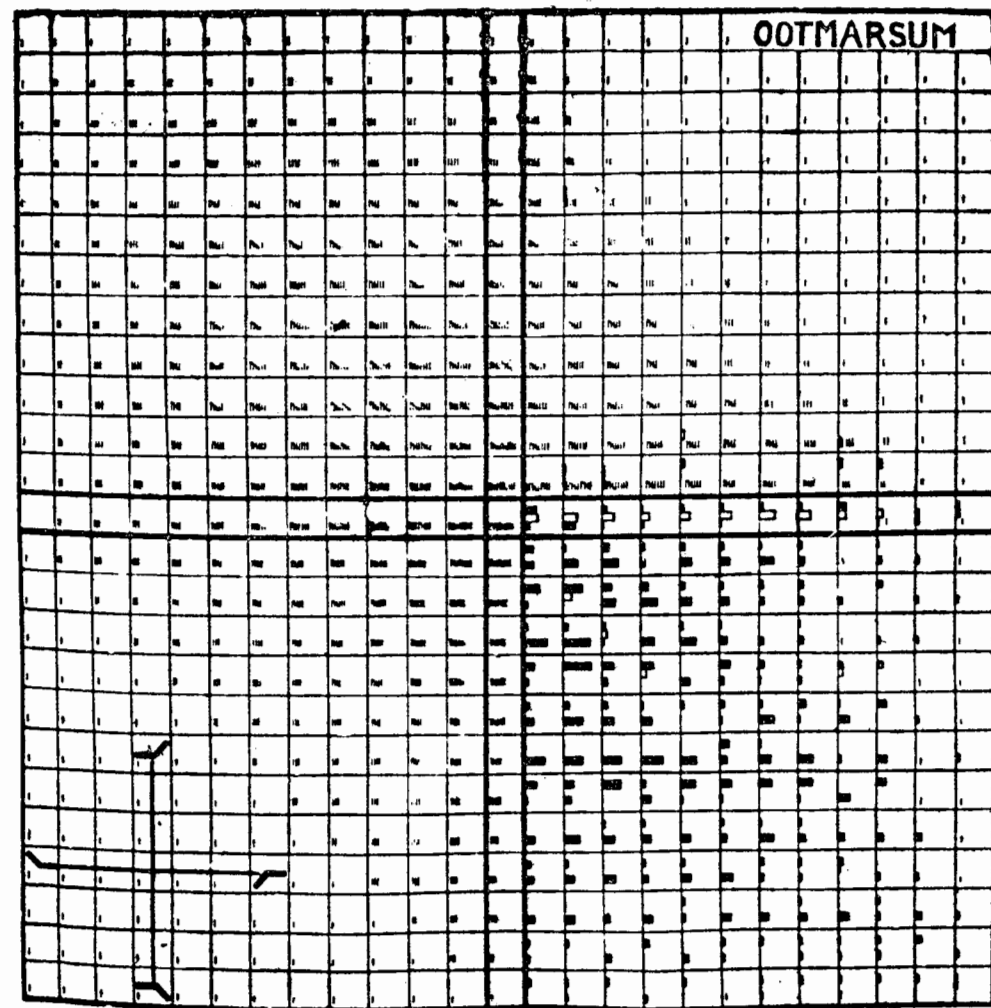


Fig. 2.

Mr De Nooyer brought all 7500 judgments into three categories: +, ±, —, though they had various forms, going from an indication of the affect by a substantive, an adjective or an adverb, to associations indicating situations, persons or animals.

The number of positive judgments for *a* was considerably larger than that for *o* and *e*. Especially cheerfulness, kindness and enthusiasm were recognized best in *a*, sadness being recognized very well in all three vowels.

As to the objective intonation curve, pitch was generally higher in the positive affects than in the negative ones, kindness and grimness forming a transition. The three negative affects were characterized by a long fall, the positive ones by a rise,

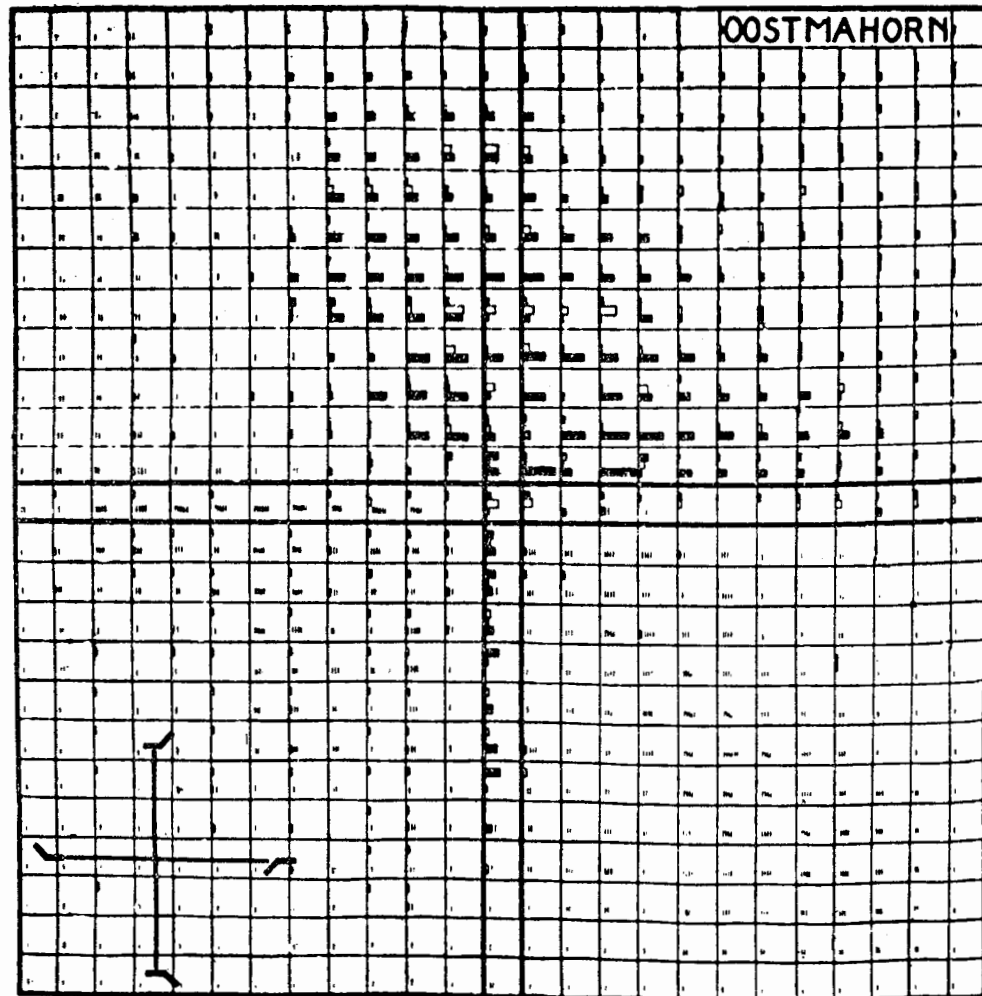


Fig. 3.

VOORLAATSTE TOON

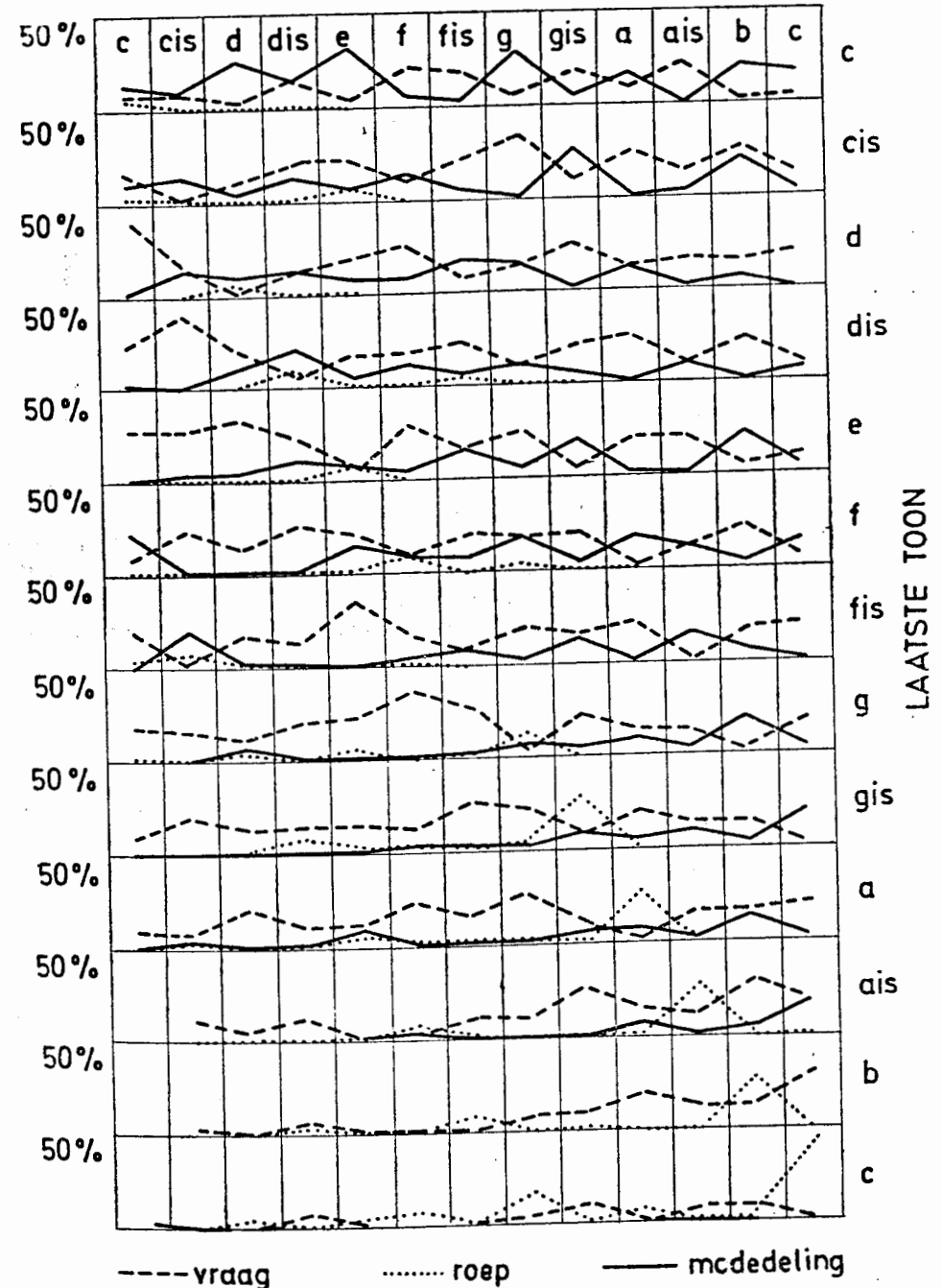


Fig. 4.

a summit and a fall. Kindness showed a final rise, demonstrating that the final part is not quite blocked for emotional information.

It will be possible by dividing the vowels into sections of about 0.5 second, to detect which part is most active.

As it is known that the subcortical region is reacting more slowly than the cortex, a longer latency and a longer excitation in the former may lead to a perfect combination of the emotional and the linguistic information.

In this connection it may be remarked that time-compression as applied by Fairbanks and his school might be dangerous by suppressing emotional information at will.

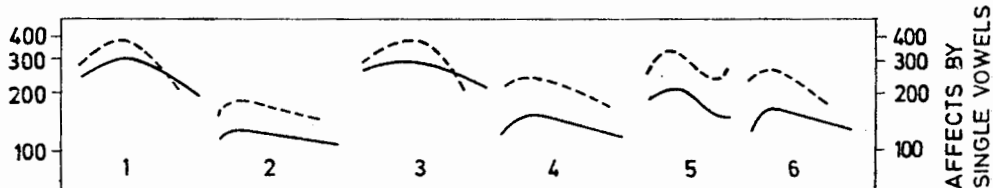


Fig. 5.

Average shape of various melody patterns in men ——— and women - - - - - The duration agrees with the average duration. The pitch of beginning, maximum and end agrees with the average values.

1. Cheerfulness; 2. Sadness; 3. Enthusiasm; 4. Disgust; 5. Kindness; 6. Grimness.

DISCUSSION

Black:

I am of course thrilled to hear the report of Professor Kaiser's pioneering work. I see in the stimulus response procedure of this work a way to seek particular "universals" in speech, ones of suprasegmental phonemes.

Fry:

We have been delighted to have the opportunity to hear Dr. Kaiser reviewing for us a greater deal of her experimental material. Her paper has suggested the idea that her results might be linked with recent experiments which have shown the dominance of one ear or the other in the reception of speech and music. It would be of very great interest if the same methods could be used in examining the distinction between affective and grammatical intonation.

MacCarthy:

Question about the technique used for producing the evidence on which the last of the speaker's slides was based.

Kaiser:

As Professor MacCarthy supposed, the last slide gave averages of the pitch records of vowels by which the subjects had tried to imitate six affects on request.

The subjects followed different ways to reach the result: some performing the affect by gestures and mimic, other limiting themselves to vocal expression.

К ВОПРОСУ О ЧЛЕНЕНИИ РЕЧИ ПРИ ВОСПРИЯТИИ

М. Г. КАСПАРОВА*

Вопрос о восприятии слов в непрерывном речевом потоке является предметом широких исследований. Эти исследования направлены на поиски тех признаков, которые описывают границы слова в речевом потоке и определяют возможность его распознавания. Известно, что эти признаки могут быть акустическими, смысловыми и языковыми. Можно предположить, что все три категории признаков сочетаются в структурном единстве и создают критерии, по которым протекает текущий и упреждающий синтез воспринимаемой речи.

Литературные данные, а также собственные наблюдения и эксперименты свидетельствуют о том, что вес каждого признака в общей сумме трех признаков неодинаков и непостоянен, являясь функцией многих факторов, которые здесь рассматриваться не будут.

Начатая нами серия опытов направлена на изучение валентности акустических характеристик речевого сигнала. В опыте, который описывается ниже, изучался не относительный вес акустического признака в системе трех признаков, а его абсолютная надежность и степень самостоятельности.

Для того, чтобы создать условия, в которых восприятие речи протекало бы только в зависимости от акустических свойств сигнала и полностью не зависело бы от смыслового содержания речи, а также от владения перцепиентом системой данного языка (имеется в виду лексическая, грамматическая и фонетическая системы) — мы отказались от методики предъявления искусственных слов и бессмысленных звуко сочетаний в чтении на родном языке, поскольку в этом случае не исключаются фонетические знания аудитора. Аудитором было предложено записать в письменном коде небольшой текст, произнесенный на незнакомом иностранном языке. Был выбран французский язык. Записывался текст удобными для аудитора знаками: русскими или латинскими буквами, или транскрипционными значками. Текст начитывался полным стилем одним диктором. В опыте приняли участие 7 испытуемых. Были получены следующие результаты.

* 1 МГПИИЯ, СССР — Москва.

С первых двух-трех прослушиваний испытуемые отказывались что-либо записывать. Приходилось делать по четыре-пять предъявлений.

В тексте, произнесенном полным стилем членение на слова внутри синтагмы оказалось достаточно удовлетворительным.

Ошибки членения на слова (т. е. определение границы слова) локализовались, в основном, на глухом согласном. *Ces cries appelaient... C'était comme un matin de chasse...* К возникновению эффекта паузы на глухом взрывном, безусловно, приводил перерыв в звучании, который возникает во время смычки.

Иллюзии паузы, кроме перерыва в глухом смычном, способствовали резкие изменения в слоговой длительности (от 50 мсек и выше) *Un coq chan|ta [fã : |tã]*.

В родном языке перерывы в звучании, изменение средней слоговой длительности приводят к эффекту членения только на стыке смысловых групп. Необходимо подчеркнуть, что эти акустические феномены обычно встречаются внутри слова, но, очевидно, происходит смысловая или грамматическая поправка, и пауза слышится не внутри слова, а после него: „Женщины были рады“.

Полученные результаты как-будто свидетельствуют о достаточной самостоятельности акустических характеристик для членения. Однако, нельзя пройти мимо других данных, которые закономерно выявились в процессе опыта.

Оказалось, что по записям auditors можно принять только 6 % переданной информации. Такое искажение возникает в результате неправильной идентификации звуков речи. Самые сильные искажения встречались при восприятии сонорных:

[rɔb] — [nɔb], [kares] — [kales], [elavɛ] — [elravɛ] и т. д.

В глухих взрывных звуки смешивались по признаку места образования: [plen] — [klen], [pɔm] — [tɔn].

В гласных ошибки возникали при идентификации звуков высокого подъема [e] воспринималось как [i] [y] и наоборот. Искажения встречались реже в носовых гласных: [ã] как [o]. Еще реже смешивались фрикативные согласные: [ʃas] как [sas]. И совсем редко искажения наблюдались при восприятии звонких согласных: [bryi] — [dui].

В результате мы пришли к выводу, что если собственно акустические признаки достаточны для перцентивного членения, т. е. определения границ слова в непрерывном потоке речи при полном стиле произнесения, то даже удовлетворительное членение недостаточно для распознавания слова. По результатам опыта трудно предположить, что распознавание слов может основываться на чисто акустических характеристиках речевых звуков. Количество переданных и принятых звуков было сохранено. Это свидетельствует в пользу пофонемного распознавания.

INTELLIGIBILITY OF SYLLABLE-TIED INTERRUPTED SPEECH

A. V. KATWIJK—J. T. HART*

Without intending to contribute to the issue of the theoretical status of the syllable, we would like to study the perceptual implications of Pilch's¹ taking the syllable as a frame within which distributional constraints can be described. Furthermore, we were struck by Huggins'² observation, that intelligibility of disrupted speech was minimal when interruptions occurred at a rate that he believed to interfere with the syllabic structure. In our experiments, using syllable-tied interruptions, to ensure that of each syllable a comparable fragment will be audible, we tried to find out which is the part of the syllable that is most vulnerable to mutilation and what is the assumed effect of distortion of the rhythmic structure of syllabic sequences. In a first experiment, described elsewhere,³ synchronization between syllables and interruptions was brought about by having the text spoken in a metrical way. Although locations of audible and suppressed fragments were not known exactly, the results strongly suggested, that the onset of the syllable, i.e. (C₁) + V, contributes more to perception than the coda, i.e. V + (C₂). As regards distributional restrictions, these very onsets constitute the distributionally free, or less predictable parts of syllables. This makes the scores, viz. 84 and 89 % in situations with audible onsets and 58 and 25 % with onsets suppressed, understandable.

Also, transgression of the distributional syllable boundary seemed to make the task more difficult for subjects.

If we now hypothesize the syllable to be a unit of intelligibility, and if we further assume that the perceptual syllable boundary coincides with the distributional one, we may expect that any addition, in the presentation of fragments, of parts of adjacent syllables, would make it less easy to allow one and only one syllable identification per presented fragment, whereas any such addition would also distort the rhythmic pattern, thus accumulating the difficulties.

The experiment to be described here is meant to test the hypotheses mentioned,

* IPO, Eindhoven, Holland.

¹ H. Pilch, *Phonemtheorie I*, Basel 1964.

² A. W. F. Huggins, Distortion of the Temporal Pattern of Speech, *JASA*, 36, 1055—1064.

³ A. van Katwijk, On Perceptual Units in Speech, *IPO Annual Progress Report 1*, 1966.

while avoiding an exaggeration of the rhythmic structure in the metrically spoken text, and securing a higher accuracy in locations of interruptions.

In preparation, a continuous text of 1100 syllables was read aloud by the experimenter at a controlled speech rate, and recorded. With a segmentator, C—V links were established and in re-recording, 120 ms before these links, impulses were put on the second track. In the experiment, these impulses triggered a time delay device, which controlled the operation of the gate with which the interruption was brought about, in six different situations, always with fixed durations of 110 ms (i.e. half the average syllable length) either in the audible, or in the suppressed parts. The six situations were (in their relation to the C—V link): 1. —35 to +75; 2. —95 to +15; 3. +30 to +140 ms audible, and 4. +15 to +125; 5. —80 to +30; 6. —35 to +75 ms suppressed. 18 Subjects were instructed to try to reconstruct the entire text, using the technique of shadowing. Training took place with 88 syllables, interrupted every 220 ms with an on-off ratio of unity. Subjects were presented with all six situations in a cyclically changing order.

The results, in percentage correct of the total number of syllables (with standard error of the mean), are: 1. 70 ± 2 , 2. 65 ± 3 , 3. 46 ± 3 , 4. 42 ± 4 , 5. 34 ± 2 , and 6. $15 \pm 2\%$.

This outcome in the first place amply confirms the main effect in the first experiment, viz. that the onsets give the greatest contribution to intelligibility. But this

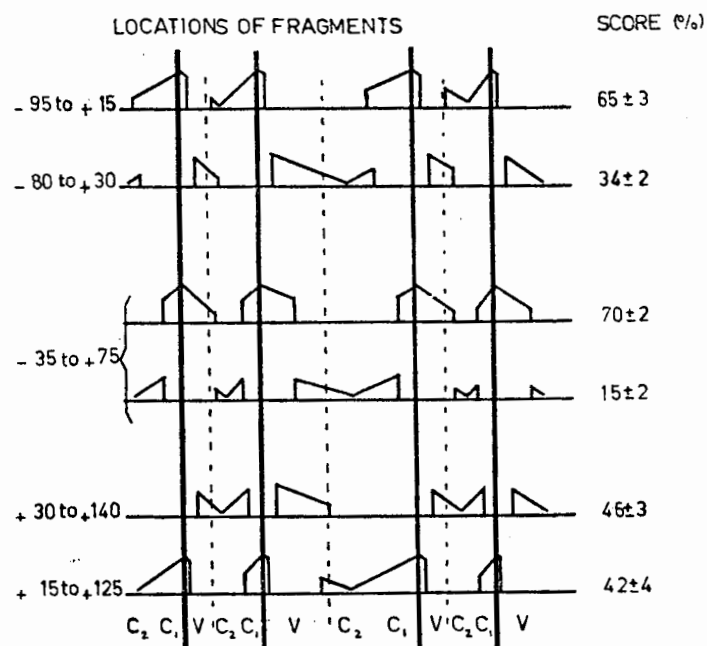


Fig. 1. Illustration of the six interruption situations used in the second experiment, for an imaginary fragment of the text. Left: opening and closing moments in relation to the C-V links: middle: audible fragments; right: mean scores and standard errors of the means.

cannot explain the difference in scores of situations 1 and 2 on the one hand and 4 on the other. We should, therefore, consider the two hypothesized conditions together, viz. a) audibility of the C—V link and b) non-interference with the rhythmic structure. It appears then, that in situations 1 and 2 both conditions are met, whereas situation 6 does not fulfil either of them. Situation 4, which allows for audibility of the C—V link, suffers from severe distortion of the rhythm, as both beginning and end of the audible part have no fixed time relation to the C—V link of the syllable from which it originates. Situation 3, which is the reverse of 4 in that it meets condition b, but not a, has about the same score as 4, showing that the two conditions approximately compensate each other. Situation 5, compared with 3, meets condition b to a limited extent in having only the beginning of the audible part related to the appropriate C—V link.

Conclusion: We feel justified in accepting the syllable as a unit of intelligibility with respect to 1. the onset, whose contribution is correlated with its distributional properties, and 2. the rhythmic pattern in syllabic sequences, whose contribution, though indirect, is well specifiable in terms of time relations with the onset.

DISCUSSION

Pulgram:

On what criteria did experimenters cut the utterance into syllables, that is, how did they know where to make the cut-off point? Is not recognizability of truncated unit due to redundancy of language rather than to distribution of phonemes? If nonsense syllables with proper Dutch phoneme distribution had been used, would not the identification possibilities have been reduced?

't Hart:

ad Pulgram: There was no need to cut the utterance into syllables; we only established C-V links.

It was our very aim to examine *where* redundancy is situated. That redundancy is not homogeneously distributed, and that there is some correlation with phoneme distribution is one of the results. Extra redundancy in our material raises the overall score, but cannot be held responsible for the differences between the various interruption situations.

ÜBER DEN URSPRUNG DER SPRACHMELODIE IN DER KLASSISCHEN UND MODERNEN DEUTSCHEN LYRIK

M. H. KAULHAUSEN

Es drängt mich, der Leitung des Kongresses für die Ehre zu danken, vor Ihnen über den Ursprung der Sprachmelodie in der klassischen und modernen deutschen Lyrik sprechen zu dürfen, ein Anliegen, das mich mehr als 35 Jahre beschäftigt hat.

Phonetik und Sprechkunde sind in ihrem Lebensgrund, der Sprache, verwandt. Aber das Verhältnis beider zur Sprache ist nicht ganz gleich. Die Phonetik interessiert das Ergebnis des Sprachaktes, der Niederschlag eines Bewußtseinsinhaltes in darstellenden Worten. An diesem mehr oder weniger rationalen Sprachelement wird die Intonation, der Verlauf der Tonhöhen und Intervalle etc., untersucht. Von den irrationalen Kräften der Sprache wird nur noch der emotionale Gefühlsausdruck im Wort bedeutsam. — Die Sprechkunde aber interessieren, als Grundlage ihrer pädagogischen Zwecke beim Dichtungssprechen, das Werden des Gebildes und die hinter den Worten stehenden Lebenskräfte, die zur sprachlichen Ausformung des Bewußtseinsinhaltes führen. Sie darf daher, neben der Verwortung und dem Ausdruck, die Mitteilungsgebärde des Gedachten und Empfundnen nicht außer Acht lassen. Das Sprachelement der Mitteilung ist in Deutschland erst zu Beginn des 20. Jhts. als die sprachwesentliche, spezifische Kontaktform des Menschen erkannt worden, die Ich und Du als „Urmodi des sprachlichen Bewußtseins“ voraussetzt. Damit hat die Sprachwissenschaft die Polarität alles Lebens, das sich zwischen Ich und Welt, der menschlichen Geist- und Triebnatur ausspannt, auch in der Sprache festgestellt.

Es ist nun gerade diese bindungsknüpfende Kraft der Mitteilung für das Werden der Melodie in der Dichtung, zu der seit Sievers auch die Klangfarbe gehört, bedeutsam. Es stimmt nicht, was gesagt wird, daß die Melodie des Gedichtes nur „in der Wortwahl, der syntaktischen Struktur und der äußern Form zu suchen“ sei, ihren Ursprung also in der vorzüglich intellektuellen Tätigkeit des Sprechers habe. Ähnliches hat auch Ed. Sievers angenommen, der mit Rutz zuerst erkannte, daß jeder Dichtung eine bestimmte, einmalige Melodie innewohnt. Beide suchten diese durch Handbewegungen und Körper Einstellungen im Sprechen zu entbinden. Beide über sahen eins: daß die Dichtungssprache keine einsame Sprache ist, sondern, wie alles lebendige Sprechen, einer Situation entspringt, in der der Sprecher das Wort an einen oder mehrere Hörer richtet. Das richtende, steuernde Element aber kann

nur die Sprachgebärde der Mitteilung sein. Denn der emotionale „Ausdruck“ ist ungesteuert und das Wort, als Wort, unbewegt. Es wird erst durch die einem Ziel zusteuende Mitteilung zum Träger einer auf- und absteigenden Melodie mit bestimmter Klangfarbe.

Dieses Ziel kann in der Klassik, die die Welt und alle Wesen der Schöpfung noch in einer mythischen, sympathischen Verbundenheit sah, ein beseeltes Objekt der Außenwelt sein, wie beispw. in Goethes „Lied an den Mond“ das Himmelsgestirn und der Fluß:

„Füllest wieder Busch und Tal
Still mit Nebelglanz.
Lösest endlich auch einmal
meine Seele ganz.“

Es kann das gesellschaftssoziale Ich eines menschlichen Hörers, und es kann das vorausgesetzte metaphysische Ich sein, das angeredet wird. Je nachdem nun ob Sinnliches oder Beseeltes in der menschlichen Umwelt angesprochen wird, entstehen aus der unterschiedlichen Mitteilungsform unterschiedliche Melodien, nach Saran und Petsch, von Dur, Moll oder einer neutralen Farbe. Ich habe daher neben der willensbetonten „Teilnahme heischenden Mitteilung“ von Ammann oder dem fordernden „Appell“ Bühlers — beide wenden sich an das Gesellschafts—Du des Hörers — zwei weitere Formen entdeckt: die Teilnahme schenkende und die Teilnahme gewährende Form. Sie kommen gelegentlich selbst im Alltagssprechen vor: im Trostspenden, Danksagen und der echten Belehrung. Sie fordern nicht und lösen keine egoistische Reaktion aus.

Hier möchte ich an zwei Beispielen zeigen, wie sich, je nach dem Ziel und der Mitteilungsform, die Melodie des Gedichtes ändert, vorausgesetzt, daß aus der Sprechsituation als der Erlebnissituation des Dichters gesprochen wird. In Goethes „Wechselied zum Tanze“ läßt der Dichter zwei Sprecher reden, den „Gleichgültigen“ und den „Zärtlichen“. Der erste fordert in einem willensbetonten Appell eine beliebige Partnerin zum Tanz auf. Der Zweite spricht die Geliebte an. Wortwahl, Syntax und Metrum sind in allen Strophen gleichartig und doch durch die verschiedene mitteilende Steuerung von verschiedener Klangfarbe, Intonation und Meloskurve und Ausgangshöhe.

- I. Komm mit, o Schöne, komm mit mir zum Tanze!
Tanzen gehöret zum festlichen Tag.
Bist du mein Schatz nicht, so kannst du es werden,
Wirst du es nimmer, so tanzen wir doch.
Komm mit, o Schöne, komm mit mir zum Tanze!
Tanzen verherrlicht den festlichen Tag.
- II. Ohne dich, Liebste, was wären die Feste?
Ohne dich, Süße, was wäre der Tanz?
Wärest Du mein Schatz nicht, so möcht' ich nicht tanzen,
Bleibst du es immer, ist Leben ein Fest.
Ohne dich, Liebste, was wären die Feste?
Ohne dich, Süße, was wäre der Tanz?

Diese Formen gerichteter Mitteilung kommen auch in der modernen Dichtung u. a. bei Rilke, Trakl und Weinheber vor, werden aber in der zeitgenössischen Lyrik immer seltener, die vorzüglich reflektierende Gedankendarstellung ist. Brechts Lyrik hat durchgehend einen lebengesättigten Durton. Bei G. Benn aber trifft man im sogenannten „absoluten“ Gedicht, das, wie er sagt (in Probleme der Lyrik) an niemand gerichtet sondern eine Montierung faszinierender Worte ist, auf melodiedürftigste Formen. Ein Beispiel aus „Fragmente“:

Stilleben

Wenn alles abgeblättert daliegt
Gedanken, Stimmungen, Duette
abgeschilfert — hautlos daliegt,
kein Stagnol — und das Abgehäutete
— alle Felle fortgeschwommen —
blutiger Bindehaut ins Stumme äugt —:
was ist das?

Es konnte natürlich in einem 10-Minutenreferat nur Wesentliches angedeutet werden. Raum und Zeit, in denen sich der Sprechakt vollzieht, mußten, neben der Vokal-Musik, dem Versmaß, der Spannung, dem Tempo, dem Kurvenverlauf außer Betracht bleiben. Wenn aber diese wenigen Hinweise zu weiteren Untersuchungen anregen sollten, wäre ihr Zweck erfüllt.

LINGUISTIC DETERMINATION OF THE INTELI GIBILITY OF THE SPEECH SOUNDS

L. KERAMITCHIEVA—S. KERAMITCHIEVSKY*

The unsatisfactory results in the phonoauditive correction with patients with impaired hearing from different linguistic areas, having in mind and using results from purely physical analysis of the phonemes isolated or in the word context, at different positions directed us to undertake a more subtle study of the problems of perception of the speech sounds from psychoacoustic and psycholinguistic points of view.

Although the Serbocroat and Macedonian languages acoustically are rather close to each other, and have seemingly close phonetic systems, the perception of the sounds in one language, especially in word and sentence context under the influence of the structural phonological elements, to the population speaking the other language is quite different and the intelligibility is accordingly more difficult. This phenomenon is more emphasized among patients with impaired hearing, therefore the problem came up in the everyday phonoauditive corrective work of our logopedic and audiological clinics.

According to the psychoacoustic analysis of the Macedonian and Serbo-croatian languages the phonemes have the following distribution:

Frequency range	Speech Sounds	
	Macedonian language	Serbo-croatian language
6400—12 800 Hz	<i>S, C</i>	<i>S</i>
4800—9600 Hz	<i>Z, DZ</i>	<i>Z, C</i>
3200—6400 Hz	<i>Č</i>	<i>I, Č</i>
2400—4800 Hz	<i>I, Š, K', NJ</i>	<i>DJ, J, NJ</i>
1600—3200 Hz	<i>E, T, J, G', N</i>	<i>E, Š, Č, T, LJ, N</i>
1200—2400 Hz	<i>P, F, C</i>	<i>Ž, DŽ, D, F, M</i>
800—1600 Hz	<i>A, K, Ž, LJ, M</i>	<i>A, R, K</i>
600—1200 Hz	<i>O, D, R, H, L</i>	<i>L, H</i>
400—800 Hz	<i>G</i>	<i>O, G</i>
300—600 Hz	<i>H, B, V</i>	<i>V</i>
200—400 Hz	<i>U</i>	<i>U, B, P</i>

(After S. Keramitchievsky and Ivo Scarich)

* Center for Phoniatriy and Audiology, Skopje-Yugoslavia.

These analyses are performed using the octave filters SIEMENS, modified in Yugoslavia, semi-octave filters ALLISON and tape recorder AMPEX. The ear was the criterion in both analyses. These two analyses were performed and verified individually and in groups in the two languages.

The procedure was that the isolatedly pronounced sound was passed through the filters, seeking the narrowest frequency band at which the sound is best heard i.e. it has the best auditive characteristic for the representatives of that language. At the same time the necessary intensity dynamic, which contributes best to the purity and perception of the sound is provided. In that way the frequency bands and the structural intensity dynamics were established at which the sounds were best heard, spoken and heard by the Macedonian speaking people i.e. Serbo-croat speaking people, regardless of the number and sex if the subjects because the preliminary investigations in this way in both languages showed that these two facts have no essential influence upon the psycholinguistic structure and perception of the sounds, although the results of our experiments in that way are such that female and child pronunciations are higher by 13—15% than the male pronunciation.

Because at these frequency ranges, with the established intensity dynamic optimal audibility of the sounds is assured and they are called optimal auditive perceptive structures. The results of these analyses have been used for several years in learning Macedonian and Serbo-croat languages by foreigners, in logopedic correction and auditive correction and rehabilitation of children and adults with hearing problems in several institutes in Yugoslavia.

The problem of linguistic determination of the intelegibility of the speech sounds appeared in the clinical practice of phonoauditive correction and rehabilitation, using the filter system with patients from the Serbo-Croat linguistic area. While instrumental phonoauditive correction, using the optimal auditive perceptive structures of Macedonian language with Macedonian speaking patients showed excellent results, whereas with the Serbo-Croat speaking patients it was not possible to achieve similar results. On the other hand Serbo-Croat children who had developed hearing loss before they learnt to speak and were rehabilitated in Macedonian institutes found it more difficult to understand even those words which are the same in both languages. Therefore comparative analyses were undertaken using two criteria for both languages: subjects with normal hearing and subjects with linear bilateral symmetric conductive hearing impairment. The experiment was performed in ten subjects of the two languages with normal and impaired hearing, using the same technique and method, which were used in establishing the frequency ranges of optimal audibility in the two languages.

The experiment aimed at establishing the difference in perception of speech sounds with regard to the acoustic and linguistic differences and hearing state. For many reasons we decided to analyse ten speech sounds of each language: five consonants *L, Č, K', G' (Ć, DJ) Ž* and five vowels — *I, E, A, O, U* under the conditions when they are:

1. Pronounced and heard by subjects of the same language, passed through their frequency ranges at the established intensity dynamic for the sounds of that language.
2. Pronounced and heard by subjects of the same language, passed through their frequency ranges, without providing the established intensity dynamic.
3. Pronounced by subjects of the same language and passed through their frequency ranges at the established intensity dynamic for that language and heard by subjects of the other language.
4. Heard by subjects of the same language and passed through their frequency ranges at the established intensity dynamic for that language, and pronounced by subjects of the other language.
5. Heard by subjects of the same language and pronounced by subjects of the other language and passed through the frequency ranges at the established intensity dynamic for the other language.
6. Pronounced by subjects of the same language, and heard by subjects of the other language and passed through the frequency ranges of the same sounds with the intensity dynamic of the other language.

The outcome of the experiment was as follows; the scoring is expressed in per cents:

Conditions	Subjects with normal hearing		Subjects with impaired hearing	
	Mac. language	S.-Cr. language	Mac. language.	S.-Cr. language
1	100	100	86	84
2	87	86	84	82
3	42	40	35	32
4	84	80	79	80
5	40	39	36	35
6	38	40	34	31

At the first condition, besides the conductive impairment of hearing of the subjects from the control group, the percentages the highest, and taking in to consideration the impairment of hearing the difference between subjects with normal and subjects with impaired hearing is minimal. As the reduction of hearing among all subjects was linear and stood at the level of 35—45 db, the emission was amplified in every case with impairment at the level of detection. The sounds the optimal frequency ranges of which were at the spectral range from 1000 Hz to 6000 Hz did not require high amplification from that one they got linearly, and the established intensity dynamic which is characteristic for the perception of those sounds by the normal ear. The optimals above and below this range required additional linear or dynamic amplification of 8—15 db. When the optimal was established in every case according to the impairment, together with the additional amplification it increased to 92 per cent.

This fact on the one hand undoubtedly shows that the optimals are standard as psychoacoustic structural form and psycholinguistic perceptive norm for every language individually. On the other hand it shows that besides the specific perception of the speech sounds by the pathological ear, which has its own system of perception is again based on the phonoauditive structure of the phonemes of the mother language.

In the modified conditions the percentage is highest when the sounds are pronounced and heard by the subjects of the same language and passed through the frequency ranges which are the optimals for the same sound of the other language. The percentage is almost the same in the two criteria. The percentage is high too, when the sounds are heard by the subjects of one language and passed through their own optimal frequency ranges with specific intensity dynamic, but pronounced by the subjects of the other language. But the percentages are much lower when these sounds are passed through their specific frequency ranges with the characteristic intensity dynamic. In that case that decreases to 40 and 30 per cent.

In the third, fifth and sixth conditions the percentage is the lowest i.e. that perception is the worst.

According to the nature of the conditions and the height of the percentage it can be easily seen that every speech sound has its defined critical frequency range and intensity dynamic, and that its psycholinguistic structure articulatively and auditively is conditioned and interdependent according to the peculiarity of the language.

Contrary to the statements that for the distinction and perception of the vowels the movement of the first and second formants is essentially important and for the consonants the width and intensity of concentrations of the acoustic energy, our previous and this experiment show it cannot be generalized, for it depends on the nature of the phoneme and circumstances. It is true that for some vowels in our language the movement of the first and second formants is more important (*O*), while for others the second and third (*U*), the third fourth and fifth as it for (*I*) etc. And the intensity of the acoustic concentration is not always important. A great number of consonants are phonologically characterized by the spectrum of the areas which physically are not the strongest. Therefore the phenomenon of perception should be sought in the phonological system of the language and not simply in the purely physical dimensions of the sounds. The structure of the physical parameters of auditive perception of the speech sounds is involved in and based on the phonological structure of the language, therefore it differs in every language even for the same or similar speech sounds in close languages. And it reflects the specific phonology and originality of the language.

Without diving into the neurophysiological, physical and psychological aspects of the problem, because in these case only the psycholinguistic phenomenon of perception of the speech sounds is taken in consideration, it can be concluded once more that the perception of the speech sounds does not depend only on the physical laws and psychophysiological conditions but it is also linguistically determined too.

DISCUSSION

Sovijärvi:

1. Did you use in all your experiments the same filter bands not varying the frequency areas? Your investigations concerning the psycho-linguistic patterns in related languages are very impressive. The continuation of your studies is important.

2. In Finnish there are 8 vowel phonemes. Therefore it would be difficult to use your filter system to distinguish the oppositions *i/y*, *e/ö*, and *ä/a*.

Keramitchieva—Keramitchievsky:

In our psycho-linguistic analysis we use a number of different filters capable of combining the frequency areas in the whole speech spectrum requiring the necessary intensity dynamics. They enable the adjustment of the electronic transmission according to the specific dimensional structure of every phoneme separately. In that case, and in connection with the first part of the question put by Mr. Sovijärvi my answer would be as follows: we use different filter bands covering all the areas capable of widening and narrowing the speech spectrum and at the same time capable of regulating (adjusting) the dimensional interrelation.

In regard to the question whether these filters can satisfy the needs for an adequate analysis of phonemes in Finnish: I am quite sure that there is no difficulty because the development of the filter technique enables us to get even more subtle results than the ordinary technique, for instance to establish opposition in the Finnish vocal system.

SOME BASIC COMPONENTS OF PERCEPTION OF SPEECH SOUNDS

S. KERAMITCHIEVSKY*

The practice in teaching deaf children to speak and understand speech, and hard of hearing children to speak and hear enables us to get a broader view of the formation classification and the interrelated attitudes of the physical and psycholinguistic distinctive features in the phonemic systems of different languages and in determining the basic components of the speech sounds i.e. the perception of speech.

The scientific separation of the problems of communication from the problems of information, although in the cybernetic phonoautitive cycle they are an indivisible whole, appears to be a strong necessity in practice because the state of these human beings deprived of the ability of normal phonoauditive communication, to receive and send information, requires a careful and complete theoretical and practical approach and solving.

While with the child with severe impairment of hearing caused during the period before the speech was learnt we can have an easier approach, because its phonoauditive „Tabula rasa“ does not lead him to the alternative choice of „yes“ or „not“, with the child with much bigger residual of hearing whose impairment had been caused after speech has been learnt, the difficulties are far greater because with the arising of the problems at the physiological level of the speech chain at the same time there happen the appropriate problems at the linguistic and physical levels i.e. there arise problems in the relation of the phonologic phenomena and the sound. And not only that. If we analyze the levels of the two halves of the speech chain separately in these cases we will realise that they are not only conditioned and interdependent, but that the problems of the one level are reflected on the other levels and by studying this phenomenon we shall clarify quite a few problems in normal conditions of communication and information.

In the first case the problem is being solved in such a way that we intervene by technical means in order to exploit the latent physiological abilities for reception of the physical stimula and transmission of the nerve impulses to the utmost in order to provide a complete sensory—motor arc. Here we use as many chanel as possible without restricting ourselves only to the auditory chanel, such as visual, tactile, kinestaetic and proprioreceptive channels.

* Center for Phoniatriy and Audiology, Skopje — Yugoslavia.

At the linguistic level, from the phono-auditive aspect maturity is conditioned by the state of the physiological level. Because the listener receives known messages in the way he had learnt them, maturity in the period of optimal biological and psychological possibilities and in changed physiological and physical conditions is almost identical, if all the necessary conditions are provided in this period of maturation for speech and hearing and if they are adjusted to the specific needs, as it is with the normal child, with the difference that he learns to distinguish the prosodic and inherent distinctive features in accordance with the state of hearing sense and ability for reception of impressions with regards to pitch, loudness and duration of the sound, phoneme or phonological structure, which are correlates of the physical dimensions of the sound. The child speaker, whose speech organs performs such movements to produce such a sound which possesses the distinctive features he is accustomed to, he controls with his own phonoauditive feed-back link and adjusts his own acoustic outcome to the distinctive features he had learnt to discriminate from complex sound waves. That means he enables himself (according to the acoustic features of his mother language) to code and recode the messages. Once the physical dimensions are changed it causes difficulties in perception so that he begins to decipher, but if the changes are greater in the physical structure, they cause changes, and in the linguistic structure he is unable to recode the message.

This rule is much more evident in the process of auditive correction and rehabilitation of children with hearing troubles which they got after they had learnt to speak. In this case the changed relation between the phonological phenomenon and the sound is more prominent because the disturbances at the physiological level of the speech chain cause impairment in the prosodic and inherent distinctive features to the extent that they distort the perception and make the speech incoherent.

While with the first technical means ought to help make maximum use of the latent auditive residual of hearing, with the second technical means should help the adjustment of the characteristics of the impressions on our senses for loudness, intensity and the subjective duration of the speech sound with the physical dimensions: frequency, intensity and time as the closest correlates of these characteristics. The quantity and quality of the distinctive features are distorted to the extent that the perception of the sounds is not possible and the speech is made incomprehensible.

This is not only because of the lack of the physical dimensions because one-sided dimensional intervention in the work with such patients did not give any results in the auditive correction and rehabilitation just because the problem was approached purely physically, tending to increase only some of the dimensions without paying attention to the dimensional structure and to the component whole at all levels of the speech chain and in accordance with the already accepted way and accustomed manner of discrimination and perception, gradually adapting distorted auditive system and enabling it in changed physiological and physical circumstances to make further use of the existing perceptual structure at the linguistic level.

It is not only necessary to satisfy the required quantitatively physical conditions

necessary for causing the stimulation to the auditive sense, but also there must be qualitative physical stimulation which through the physiological and linguistic level would provide an adequate response to the messages. It does not mean that it is necessary to provide absolutely identical structure, because it is impossible, but to provide modified structure being so close to the one he is accustomed to so that he is able to respond to the distinctive features and types of structural elements which participate in the transmission of the information according to the principles of the code. It is those elements which every member of a linguistic community is taught to manipulate and which constitute the linguistic specifics, because phonemes in one language are not the speech sounds themselves but the elements of those sounds which are in a certain way realised in the appropriate contexts and which these children and adults learnt how to pronounce and recognize in different combinations of phonemic sequences.

That means the perception of speech sounds i.e. phonemes is very complex and conditioned by the presence of the physical, physiological, psychological and linguistic components which constitute its whole and specifics and makes the process of coding and recoding of the information possible. Therefore it should be observed and studied as a whole, which cannot be easily simplified as is often done.

DISCUSSION

De Grève:

Approving, of course, Mr. Keramitchievski's distinction between a structural dimension and a mere physical dimension, which he opposes to each other, I would like to know how he operates to fix the structural dimension. On the other hand, I would like to know whether Mr. Keramitchievski distinguishes between different kinds of non-normal hearing, being the distinction between a normal ear and a non-normal ear.

Keramitchievsky:

The psycho-linguistic structure of the phonemes in many languages in the world is determined in such a way that as a basic criterion of discrimination, identification and verification of the information is not considered only the ear as a constant organ which reacts physiologically to a fixed physical dimension or dimensional context, but also the hearing which at the psycho-linguistic level of the speech range represents a condition and means for linguistic communication within the same linguistic area. This means that in order to determine the structural basis of the phonemes in a given language the sound is not taken either as a simple addition of the physical dimensions nor simply the articulated sound which in order to be a real sound ought to possess linguistic structure, but also the sound (phoneme) itself, belonging to the phonetic system of a given language and in which the dimensional parameters are specific; and their interrelation forms the distinctive phonological structure, the only difference being between the mere physical and structural dimensions. In connection with the second part of Mr. Grève's question I can only say that the normal ear is normal with the subjects of both linguistic areas, while the same type of impairment was taken in the two groups.

SOME PHYSIOLOGICAL SPEECH PROBLEMS IN HIPERBARIA

STANISLAW KLAJMAN

Economical conditions and the development of our civilisation draw our attention to scientific researches of the undersea world and to its industrial exploitation. Apart from many technical and physiological problems one of the most virtual is to assure the diver a good verbal communication during his work in gas-hiperbaria. The researches of Hollywell and Harvey (2), Mac Lean (5) and Golden (1) concerned only acoustical analysis of divers speech in helium-oxygen mixtures. The aim of this report is to demonstrate the results of our investigations on speech in air-hiperbaria. Investigations performed on the physiology, pathology and acoustics of a divers voice in air overpressure we described in a separate paper (3). Common rules formulated in our works could be extrapolated to different gas mixtures, taking into account their specifics. Specifics of various languages must be also considered.

MATERIAL AND METHODS

The investigations were performed with 20 young, healthy, professional Polish sea-divers, who are phoniatically correct. The text used for the investigations was: "głosem dźwięcznym" which means: "with sonorous voice". The divers staying in the pressure chamber read the text (recorded on the magnetic tape AGFA) at pressure levels: normal, 1, 2, 3, 4, 5 and 6 atm. gage pressure. They were trained to read the sentence properly and to maintain a steady level of loudness, which was also checked on the level meter. The tape recorder was provided with capacity Brüel-Kjaer type 4133 microphone. The frequency range of the tape recorder was from 20 Hz to 18 kHz. The pressure chamber had a good acoustical characteristics and the reverberation measurements, carried out at various pressure levels, enabled us to eliminate its influence on the acoustic structure of recorded speech. Acoustical analysis of the text was performed using the Key Electric "Sonograph". May I express, just now, my heartiest thanks to Prof. Dr. Sc. M. Seeman, Head of the Phoniatic Clinic in Prague and m.d. A. Novák for their willingness and valuable advices, as well as for enabling me to use the instrument.

RESULTS

The tested text consists of vowels and consonants, mouth and nasal sounds. The sonagrams were evaluated on ground of direct observations and measurements of the formant bands using a photoelectric transducer. The frequency bands of sonograph filters were 300 Hz. The acoustic structure of the tested text and the time parameters at different pressure levels were compared using the "t" Student's test.

Depending on the pressure in the chamber we stated: 1. increase, damping or decay of some frequencies, 2. translocation of the formant bands into higher frequencies, 3. occurrence of a formant band in the range at about 4—5 kHz and increase of this range of nasal sounds. It is bound with the insufficiency of the soft palate, which causes the increase of the nasal resonance, 4. shape changes of some formant bands, 5. decay of slit sounds such as s, c, cz. It involves lacks in the text and changes its logical structure, 6. elongation of the phonation time of the text from 1,55 sec. ($\delta = 0,09$) average to 2,05 sec. ($\delta = 0,13$) average. All these changes were bounded with the mechanical influence of air-hyperbaria on the voice organ and the narcotic effect of air-hyperbaria on the central nervous system of the diver and on his hearing process (4). The stated changes were statistically significant ($p < 0,01$). A sample of an acoustical analysis of the text read out by one diver in normal air pressure and 6 atm. gage pressure is given below on fig. 1 and 2.

The text: "głosem dźwięczącym" at normal air pressure, according to Boltzmann's formula (6), has the entropy of 1,2041 d.u. average per stimulus. But at 6 atm. gage pressure the average information content, according to Shannon's formula (6) diminishes to 1,079 d.u. average per stimulus. The average information content of the text read out by 20 divers, with regard to sounds, lowered statistically significantly ($p > 0,01$). The additional factors which increase the desinformation is the change of the timbre and the pitch of the sounds and a lengthening of the phonation time. Generally the speech of divers in air-hyperbaria is slow, high pitched and nasal. The changes in the information content of speech in air hyperbaria, in which also other factors affect the diver (e.g. emotion, the influence of breathing noises etc.), lead to desinformation and to errors in submarine operations which could end tragically. Because in few cases the intelligibility of speech was good, it seems that suitable persons for proper phonation in hyperbaria should be selected using tests. By training of the vocal organ some divers could be also adapted to such conditions.

CONCLUSIONS

1. The specific influence of air-hyperbaria on the central speech analyser and the peripheral speech organ involves changes in the acoustical structure of divers' speech and is the reason for disturbances in its information content.
2. Apart from the vocoder correction of speech in hyperbaria, phoniatric training should be introduced to adapt the speech organ of divers to work-conditions in hyperbaria.

REFERENCES

1. Golden R. M.: Improving naturalness and intelligibility of helium oxygen speech, using vocoder techniques. *JASA* 1966, 40, 621—624.
2. Holywell K., Harvey G.: Helium speech. *JASA* 1964, 36, 210—211.
3. Klajman S.: Wpływ hiperbarii na głos ludzki. WSMW, Gdynia 1966.
4. Klajman S., Ruszel J., Torbus J.: Przyczynek do badania krzywej progowej słuchu w warunkach wysokiego ciśnienia u nurków. *Lek. Wojsk.* 1960, 6, 578—581.
5. Mac Lean D. J.: Analysis of speech in a helium-oxygen mixture under pressure. *JASA* 1966, 40, 625—627.
6. Wartak J.: Metody cybernetyczne w biologii i medycynie. PWN, Warszawa 1966.

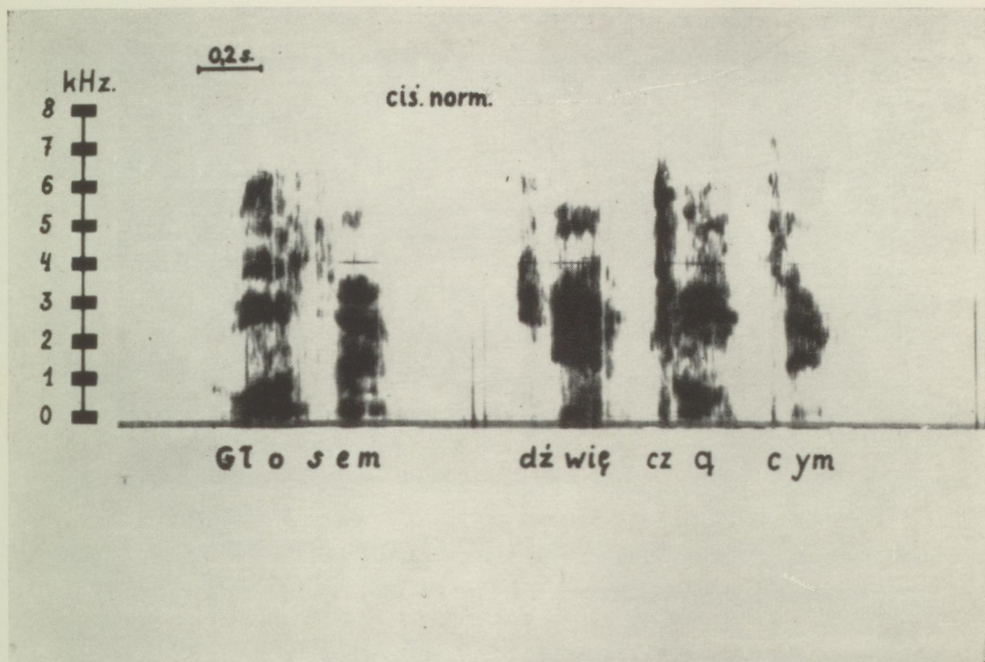


Fig. 1.

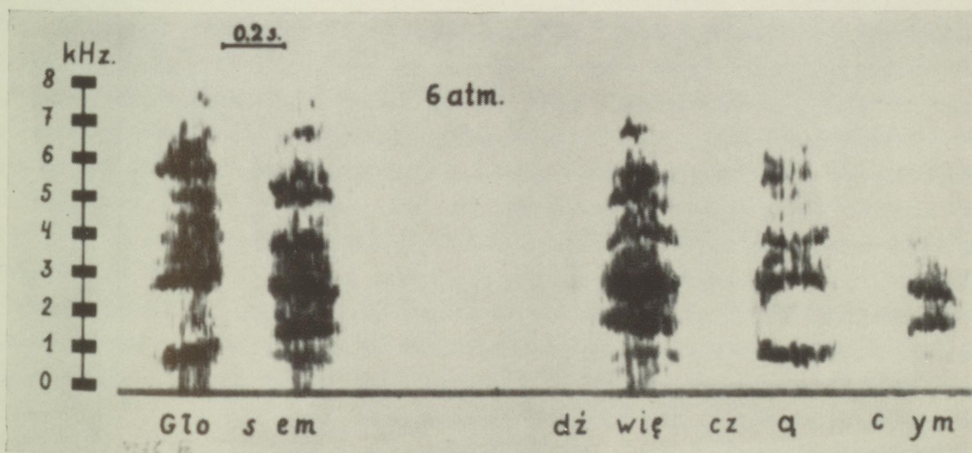


Fig. 2.

W. A. KOCH

0. Folgende Leitfragen mögen die Überlegungen bestimmen:

- (1) Gibt es spezifische phonische Einheiten, die in *spezifischer* Weise zur Textstrukturierung beitragen?
- (2) Welche Textstrukturen werden gegebenenfalls durch solche phonischen Einheiten differenziert?
- (3) Welche Modifizierung bezüglich der üblicheren Vorstellungen vom *Phonem-inventar* bietet sich u. U. nach einer Konzeption der Textstrukturierung an?

1. Bevor die Frage nach dem spezifischen Verhältnis zwischen Phonem und Text angegangen wird, soll der theoretische Status dieser beiden Größen kurz erörtert werden. Der Text soll nicht etwa ein mehr oder weniger *amorpher* Sammelplatz aller derjenigen Einheiten sein, die mit den akzeptierten Mitteln einer Syntax, Morphologie o. ä. nicht mehr beschrieben werden können; er soll als Ebene (Planum) *sui generis* betrachtet werden; er wird Eigenschaften und Strukturierungen aufweisen, die mit denen des Morphems und des Satzes vergleichbar sind. Wir bedienen uns im folgenden eines besonderen Planifizierungsmodells. Danach erhalten wir eine Folge vergleichbarer Plana: Das Morphem, das Logem (grob vergleichbar mit der Größe „Wort“), das Syntaktem (vergleichbar mit „Satz“) und das *Textem*. Vom *Textem* ab erhalten wir das Satellitplanum „Situationem“. Eine der Strukturen *innerhalb* des *Si* soll der Fokus des Textteilnehmers sein. Der Fokusmodus, der die zunächst zu besprechenden Größen bestimmt, soll der *informationelle* genannt werden. Es kann hier nicht auf die besonderen Einzelheiten der Systematik eines solchen Modells eingegangen werden.¹ Wichtig scheint in diesem Zusammenhang vor allem der Gedanke, daß das Phonem nicht in die Reihe der Planifizierung (Reihe 1) gehört, u. a. auf Grund der Vorstellung, daß die Funktion mancher Phoneme so dargestellt werden kann (oder soll), daß sie erst auf höheren Plana einsetzt. So durchläuft etwa das Phonem [ɰ] nicht die Plana Morphem-Logem-Syntaktem (wie etwa ein Phonem wie [l] oder [i]), sondern setzt in seiner Funktion erst auf dem Planum „Syntaktem“ ein. Die Phoneme (in den Reihen 2, 3, 4) sind demnach von der Planifizierung (Reihe 1)

¹ Vgl. Koch, W. A., *Taxologie des Englischen* (in Vorbereitung).

getrennt und können sich direkt auf spezifische Plana beziehen (wie das Schnalzphonem [ʔ], das direkt ein Syntakt oder ein Text der „Mißbilligung“ o. ä. manifestiert). Die Phoneme sollen ferner danach unterschieden sein, ob sie zur Manifestierung des „Kerns“ eines Planums beitragen (wie z. B. — in der Reihe 2 — [l], [t] und [ŋ], die das NE Morphem <ling> manifestieren können), ob sie Strukturen *innerhalb* eines Planums manifestieren (wie z. B. ['] und [ˈ] — in der Reihe 3 —, welche bestimmte Logem- oder Logemfolgestrukturen differenzieren, wie etwa <dancing-girl> und <dancing girl> —, oder ob sie Grenzsignale der Plana sind (wie z. B. das *junction*-Phonem für das Morphem oder [#] für das Syntakt — Reihe 4 —). Diejenigen phonischen Einheiten, die uns hier im besonderen interessieren, siedeln sich also in der Reihe 3 unterhalb des Textems an: Wir erwarten, daß Einheiten wie „spezifische *stress*-Folge“, „spezifische Tonkurven“, „schnelleres ['] oder langsames ['] Sprechtempo“, „Tremolo“ [TR] u. a. in spezifischer Weise Textstrukturen manifestieren.

2. Das Inventar der für die Textstrukturierung in Betracht kommenden Phoneme erscheint vorläufig als offen. Die tatsächliche Relevanz mag sich durch Kommutationstests o. ä. Prozeduren erweisen. Diese wiederum werden dadurch bestimmt sein, welche und wieviele Textstrukturen wir *explizit* registrieren. Betrachten wir die Opposition ['] : [ˈ] auf den Graphen 2.1 und 2.2. Sie wird ausgewiesen durch die Verteilung ['] für den *Topik*, [ˈ] für den *Komment*.² In 2.1 ist also <glass> der *Topik*, in 2.2 ist es <elastic>. Wir betrachten die *Topik*-*Komment*-Dichotomie als die niedrigste der Textstrukturen. Gehen wir zur Opposition [ˈˈˈ] gegenüber („Absenz von Tempowechsel“) auf den Graphen 2.3 und 2.4 über, so kann <to be or not to be> in 2.3 gegenüber dem Rest <that ist the question> als *Kommentarsyntakt* gelten, während die vergleichbaren Einheiten in 2.4 ein homogenes Syntakt bilden. Betrachten wir die Opposition in 2.5 und 2.6 Tonkurven zeichnen in besonderem Maße *Komments* aus. Die „ungewöhnlichen“ Kurven in 2.5 suggerieren demnach einen besonderen *Komment*-Charakter der betreffenden Einheiten, welcher in 2.6 in normaler Weise durch die normale syntaktische Struktur nicht vorhanden ist bzw. auf <wood> beschränkt ist. Man kann auch sagen, die Einheiten in 2.5 werden über die syntaktische Vorinterpretation hinaus dem topikal-fokalen Verkettungsfokus in besonders starker Weise angeboten. So macht in 2.5 das syntaktisch als „weniger wichtig“ (d. h. hier „attributiv“) ausgewiesene <dark> durch die Tonkurve besonders auf sich aufmerksam. Eine äquivalente Prozedur, die syntaktisch niedriger eingestuften Einheiten für die *Topik*-*Komment*-Verkettung stärker anzubieten, bestünde darin, das involvierte Syntakt zu *devolvieren*, d. h. zur „Parataxe“ zu bringen, etwa: *There was a girl. She was called Red Riding-Hood. She was small. She went through a wood. The wood was dark etc.*³ Der „unsystematisch“ erscheinende Märchen-

² Vgl. Firbas, Jan, 1966, Non-Thematic Subjects in Contemporary English, in: *Travaux linguistiques de Prague* 2 (1966).

³ Vgl. Koch, W. A., Transformationsrelata und ihre Vorkommensmatrix (in Vorbereitung).

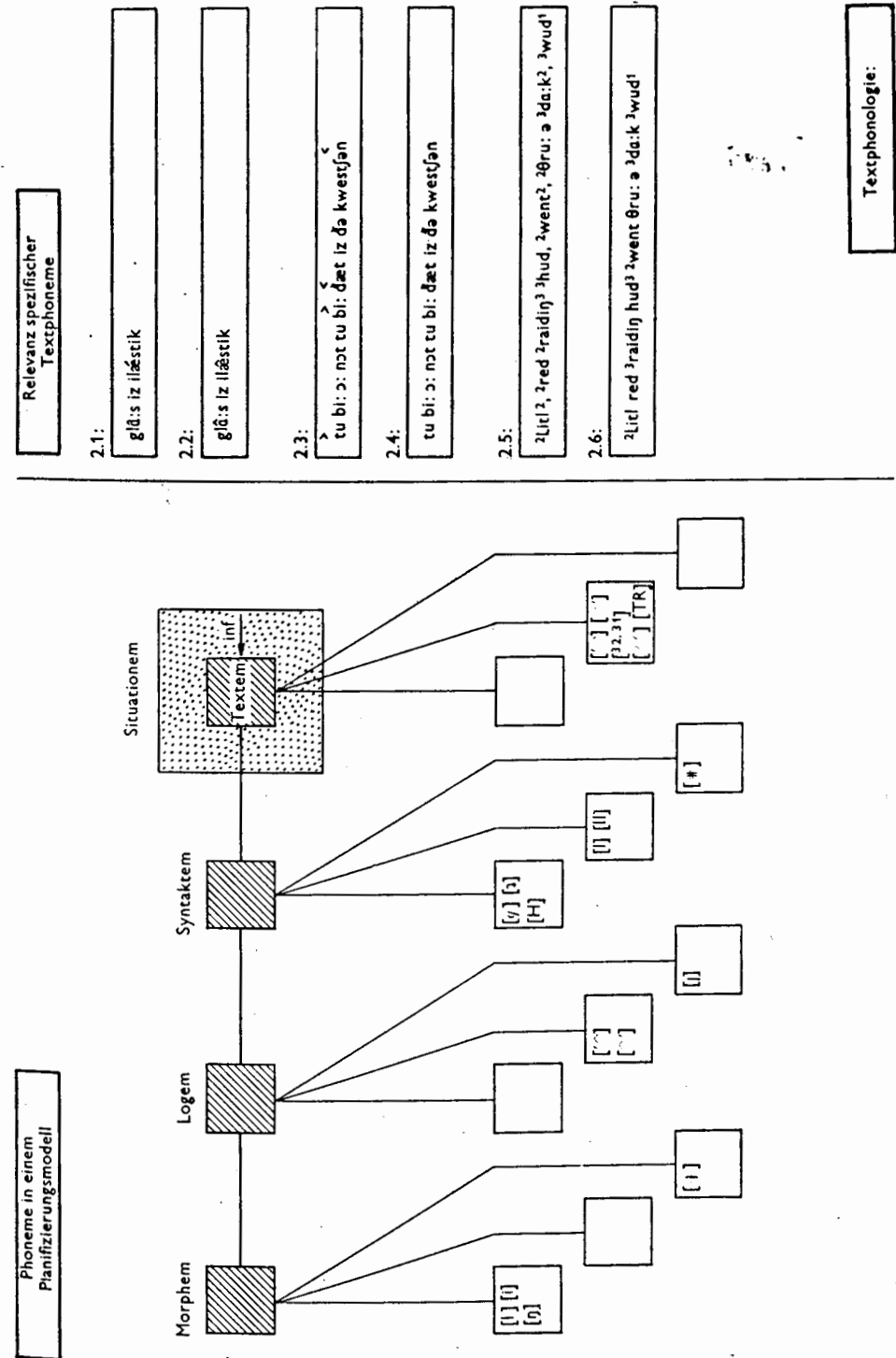


Abb. 1.

Beispiele verschiedener Textstrukturen

4.1: nou ðen ðaisélf | prizju:m not góð te skæn # ða própa stúdi av mænkaínd iz mæn #

4.2: nou ðen ðaisélf prizju:m not góð te skæn ða própa stúdi av mænkaínd iz mæn

4.3: ða windz bæstina:dou wipt on ða kæl ikou skin av ða mækeru:n
TR—TR

4.4: ða windz bæstina:dou wipt on ða kæl ikou skin av ða mækeru:n

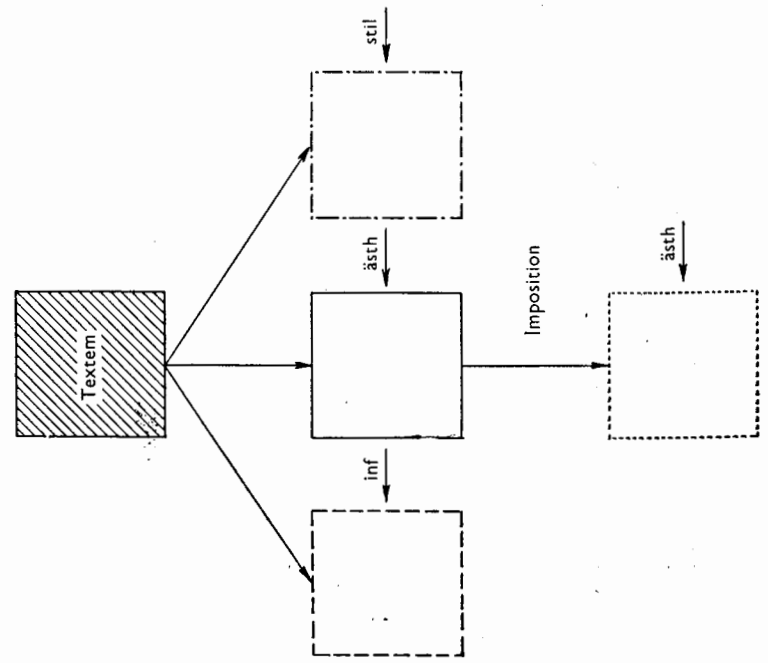
4.5: ʔmaŋʃ ʔbjutʃ ʔiz ʔ les ʔ ʔðæn ʔ ʔdʒ ʔ ʔfeis ʔ ʔyʔ # ʔmaɪ ʔ ʔda:k ʔ ʔhjarou ʔ #

4.6: ʔmaŋʃ ʔbjutʃ ʔiz les ðæn ða ʔfeis ʔ av maɪ ʔda:s:k ʔhjarou ʔ #

Textphonologie:

Abb. 2.

Phoneme verschiedener Modi von Textstrukturen



ton wäre demnach durchaus mit expliziten syntaktischen und textuellen Strukturen zu verbinden.

3. Wir erwarteten, daß sich bei gesteigerter Explizierung der Textstrukturen das Inventar der relevant werdenden Phoneme vergrößert. Bisher sind nur die vom Fokusmodus der Information erzeugten Strukturen gestreift worden. Der auf der Höhe des Planums Textem erscheinende Satellit *Situationem* enthält jedoch noch andere Fokusmodi, z. B. den ästhetischen und den stilistischen. Diese erzeugen Strukturen *sui generis*. $\leftarrow_{\text{ästh}}$ ist auf nicht-triviale Rekurrenzen gerichtet, $\rightarrow_{\text{stil}}$ auf Durchbrechungen spezifischer Erwartungen. Die unter $\leftarrow_{\text{ästh}}$ ausgenutzten \rightarrow_{inf} -Strukturen können nun in besonderen Fällen so angeordnet werden, wie sie unter reinem \leftarrow_{inf} -Fokus nicht zulässig wären. In diesem Fall sprechen wir von *Imposition*.⁴ — Es könnte nun überprüft werden, welche phonische Einheiten bei Kommuntation (bzw. Absenz) diese zusätzlichen Strukturen vernichten würden.

4. Auf dem Graph 4.2 (gegenüber 4.1) würde eine Absenz der *streß-* und *Pausenphone* nicht nur \leftarrow_{inf} -Strukturen, sondern in noch radikalerer Weise $\leftarrow_{\text{ästh}}$ -Strukturen auslöschen. In 4.4 (gegenüber 4.3) würde die Absenz der Folge „Crescendo//Decrescendo“ sowie des Tremolos entscheidende stilistische Strukturen des Gedichts von E. Sitwell zerstören. Diese zusätzliche phonische Kennzeichnung der \leftarrow_{inf} -Einheit *wind* durchbricht unsere normale Erwartung und weist sich so als *Stilistikum* aus. Die in 4.6 gegebene Lesung eines Gedichts von José Garcia Villa wurde (auch im weiteren Zusammenhang des Gesamttextes) den Möglichkeiten der metrischen ($\leftarrow_{\text{ästh}}$) Strukturen starken Abbruch tun. Die in 4.5 gegebene *Imposition* (die vom Autor graphisch suggeriert wird) ergibt eine stärkere ästhetische Sättigung und weist in *diesem Sinne* etwa die Tonkurvenfolge in 4.5 gegenüber in 4,6 als *phonematisch relevant* aus.

5. Das Phoneminventar einer Sprache bleibt folglich in dem beschriebenen Sinne solange offen, als man nicht in der Lage ist, eine *endliche* Anzahl von Textstrukturen zu explizieren. Es gibt natürlich weiterhin relativ geschlossene Untermengen. Die übrigen — weniger geschlossenen — Untermengen könnte man freilich als nicht-sprachenspezifisch der Paralinguistik⁵ zuweisen. Vielleicht würde man sich jedoch dann eines entscheidenden Stimulus zur Etablierung von Textstrukturen als Strukturen *sui generis* begeben.

⁴ Vgl. Koch, W. A., 1966, *Recurrence and a Three-Modal Approach to Poetry*, The Hague: Mouton.

⁵ Vgl. Pittenger, R. E., Hockett, C. F., Danehy, J. J., 1960, *The First Five Minutes — A Sample of Microscopic Interview Analysis*, Ithaca, N. Y.: Matineau.

VOWEL, CONSONANT AND SYLLABLE RE-EXAMINED

KLAUS KOHLER*

Hosts of phoneticians have attempted to define, in phonetic terms, the fundamental division which we recognise between *vowels* and *consonants*. K. L. Pike, in the 6th chapter (Classification Criteria) of his *Phonetics* (Ann Arbor 1944, p. 66ff) examines these various definitions critically, and replaces them by a different one of his own. "All vocoids are simultaneously *vowels* at the time that they are functioning as syllable crests... All contoids while functioning as non-syllabics are *consonants*." (ibid. p. 145) He introduces a functional view into the discussion and contrasts his purely phonetic terms *vocoid* and *contoid* with the traditional terms *vowel* and *consonant*, which in his definition now refer to the role they play in the syllable.

J. D. O'Connor and J.L.M. Trim ('Vowel, Consonant and Syllable—A Phonological Definition'. *Word* 9 (1953), 103ff) as well as G. F. Arnold ('A Phonological Approach to Vowel, Consonant and Syllable in Modern French; *Lingua* 5 (1955—6), 253ff) went one step further and, although still aiming at a phonological rather than a phonetic solution, freed themselves from the syllable as a prerequisite for such a definition. Instead they investigated the distributional properties of phonemes in English and French words and found that there are characteristic groupings in both languages the basic one of which corresponds closely to the traditional vowel-consonant distinction.

But these conclusions which were reached after an empirical examination of a very limited corpus of data is only valid for English and French, whereas the general assumption is that we need vowels and consonants in the description of all languages. As we cannot hope to investigate some 3000 languages spoken on this planet as to their phoneme distribution in words in the near future only a general theoretical reflection about the characteristics of natural languages will bring us nearer the goal.

As a child is able to "extract", in an extraordinarily short time, the structural laws of a language from the corpus presented to him the relevant data must be of such a size that the child can really be confronted with it. If now in a language phonemes can be combined in any possible way to form formatives the number of different combinations (even without repetitions of phonemes) takes on unmanageable values once the

* Institut für Phonetik und Kommunikationsforschung, Bonn, W. Germany.

set of phonemes goes beyond a certain limit. If there are 7 phonemes the result is 13692, which means that the child must come across at least 13692 formatives in the first six or so years of his life to discover the combinatorial laws. This is impossible. Therefore a language must possess less than 7 phonemes if all the combinations are to be possible, but a language of this very simple kind is not known. Thus the simplest combinatorial rule finds no application in natural languages because it would in fact make the pattern extremely complex and phonologically irregular. It can be proved mathematically that the simplest and most efficient set-up is reached if there are always two groups of phonemes such that the combinations *between* the groups are very little restricted, whereas fairly strict and regular limitations are imposed on combinatorial possibilities *within* the groups, the *degree* of restriction depending on the particular language in both cases, but being always lower in the former than in the latter. We can now call that group *consonants* that contains the stops, the other *vowels*, and have thus arrived at a substantive universal for all natural languages.

The next question is whether something similar holds for the *syllable*. Like the previous two terms it has been given a great many different definitions, phonetic and phonological. The surveys by B. Hála (La syllabe, sa nature, son origine et ses transformations, *Orbis* 10, 1961, 69ff), G. Laziczus (Lehrbuch der Phonetik, Berlin, 1961, 156ff) and A. Rosetti (Sur la théorie de la syllabe, s'Gravenhage 1956) illustrate this diversity of theories. Like the other two terms the syllable has also been regarded as a substantive *phonological* universal, but it can be demonstrated that the syllable is either an *unnecessary* concept, because the division of the speech chain into such units is known for other reasons, or an *impossible* one, as any division would be arbitrary, or even a *harmful* one, because it obscures underlying structures.

If the syllable has any real status in phonology its boundaries must be discernible. A grammatical formative in any language can be rewritten as a sequence of elements like

$$GF \rightarrow C_1VCVC\dots CVC_2$$

where C_1 stands for any non-arbitrary pre-vocalic consonant (cluster), C_2 for any non-arbitrary post-vocalic consonant (cluster) and C for any consonant (cluster) between vowels. Any part to the right of the arrow may be zero.

C can now be rewritten as

$$C \rightarrow C_1, C_2, C_3, C_2C_1$$

($C \rightarrow C_1C_2$ is impossible by definition).

If $C_1 \neq C_2$ and if $C \rightarrow C_1, C_2, C_2C_1$, a non-arbitrary syllable division is possible but implicit in the definitions for C_1 and C_2 : the syllable is therefore an *unnecessary* concept.

If some $C_1 = C_2$ there are cases where a syllable division is not determinable, and the syllable is therefore an *impossible* concept. The same applies to $C \rightarrow C_3$, i.e. a con-

sonant (cluster) that can only occur intervocalically and can, therefore, not be uniquely determined as pre- or post-vocalic.

If in any language the division into syllables obscures underlying phonological structures the concept of the syllable is *harmful*. English provides an example.

There intervocalic consonant sequences are of any degree of frequency only if they are either pre- *and* postvocalic at the same time (*butter, master*) or if one consonant is ambivalent, i.e. is part of a C_1 *and* part of a C_2 simultaneously (*mattress, belfry*). If there is clear separation between C_1 and C_2 the sequences are rare (*aesthetic, Norway*). The underlying structure of di-vocalic words in English therefore depends on syllabic indeterminacy; the introduction of the syllable into the analysis can only obscure this fact.

In languages in which a non-arbitrary syllable-division is always possible because it is implicit in the consonant clustering and in which this division does not obscure the underlying structure it is still possible to use the syllable as a unit for quick reference, although it is not necessary because the syllable is then not an independent entity. In all other languages the syllable has no place in the phonology, and it is, consequently, not a phonological universal.

PERZEPTIONSUNTERSUCHUNGEN ZUR RELEVANZ DES COUP DE GLOTTE IM DEUTSCHEN

EVA-MARIA KRECH*

Der coup de glotte wird im Deutschen als eine Form des Stimmeinsatzes bei Vokalen im Stamm- und Präfixanlaut verwendet. Er findet sich in diesen Positionen etwa dreimal so häufig wie der weiche Einsatz und ist daher als eine charakteristische Erscheinung der deutschen Hochlautung anzusehen. Außerdem dient der Glottisschlag als Grenzsignal, da er ausschließlich bei der Randstellung der Vokale vorkommt, während der weiche Einsatz auch im Morpheminnern erscheint (2).

Angesichts dieser Tatsache erhebt sich u. a. die Frage, ob die richtige Perzeption bestimmter Bedeutungseinheiten vom Vorhandensein des coup de glotte abhängt oder beeinflußt wird.¹ Als Beitrag zur Beantwortung dienten verschiedene Hörtests: 20 Gegensatzpaare vom Typ *mitteilen* — *mitteilen*, *Schiffart* — *Schiffahrt* wurden von 2 Sprechern jeweils im Satzzusammenhang auf Tonband aufgesprochen und in 4 verschiedenen Testreihen, vom Kontext getrennt, insgesamt 57 Hörern zur Perzeption vorgeführt. Während sich Sprecher A seiner üblichen Lautungsart bediente, hatte Sprecher B gezielt ausgewählte Varianten zu realisieren:

1. Formen, bei denen der betreffende vokalische Anlaut zum vorangehenden Laut in open juncture steht und mit Glottisschlag einsetzt,
2. Formen mit open juncture ohne coup de glotte und
3. die entsprechenden Formen mit close juncture.²

Weitere Beispiele kamen durch gezielte Akzentveränderungen hinzu.

Insgesamt standen zur Auswertung 7 858 Antworten zur Verfügung. In 1032 Fällen, das sind 13,1 %, urteilten die Perzipienten falsch. Diese Fehlinterpretationen verteilten sich nahezu gleichmäßig auf sämtliche Hörer, waren aber auf bestimmte Varianten konzentriert:

* Institut für Sprechkunde und Phonetische Sammlung der Martin-Luther-Universität Halle-Wittenberg

¹ Bei Meyer—Eppler (4) heißt es: „/?/ ist das Phonem, das beispielsweise die Wörter *verreisen* . . . und *vereisen* . . . zu unterscheiden gestattet“. — Hinweise auf eine gelegentliche bedeutungsdifferenzierende Wirkung von [ʔ] finden sich u. a. auch bei Korlén/Malmberg (1).

² Zur Frage nach der Verwendung von open und close juncture im Deutschen vgl. bes. Moulton (5,6).

Tab. I. Perzeption von Formen mit open u. close juncture.

	Open juncture				Close juncture	
	-ʔV	-ʔV	-V	-V	V	V
Durchschnitt in %	100	94,4	91,5	76,8 13,7	98,5	94,4

Die Formen, bei denen open juncture vorlag, der Anlautvokal betont war und mit [ʔ] einsetzte, wurden in jedem Fall einwandfrei identifiziert (z.B. *mein* 'ʔ *Eid*, *Schiff* 'ʔ *art*). Fehlte bei sonst gleichbleibenden Verhältnissen der Akzent, dann gab es im Durchschnitt noch 94,4% richtige Urteile. War jedoch umgekehrt die Einsatzsilbe akzentuiert, ohne daß dabei ein [ʔ] realisiert wurde, so ließ die Sicherheit in der Bestimmung weiterhin etwas nach; 91,5% der Fälle wurden hier im Durchschnitt erkannt. Die größte Schwierigkeit für die richtige Identifikation bereiteten die Formen, bei denen ein unbetonter Vokal zum vorangehenden Laut in open juncture stand, aber ohne [ʔ] einsetzte. Hier stimmte das Urteil im Durchschnitt nur in 76,8% der Fälle. Außerdem gehören in diese Gruppe die gesondert gefaßten Beispiele wie *würd' ich*, *schneid' ich*, die im Durchschnitt überhaupt nur in 13,7% der Fälle richtig erkannt wurden. — Bei den Formen mit close juncture ließen sich ebenfalls die Testwörter, bei denen die betreffende Silbe akzentuiert war, besser bestimmen, als die unbetonten Formen.

Zusammengefaßte Ergebnisse einer Einzelanalyse der Varianten:³ In allen Fällen, in denen der anlautende Vokal mit einem vorangehenden R, einem Verschuß- oder einem Englaut in open juncture stand, war der Gebrauch des [ʔ] zur Identifikation nicht notwendig, ausgenommen die bereits erwähnten Beispiele *würd' ich*, *würdig* usw. In Wortpaaren wie *der Ost* — *der Rost* erwies sich das Fehlen oder Vorhandensein eines anlautenden R-Allophons für die Perzeption als relevant. Lag open juncture vor, waren vorangehende Enge- und Verschußlaute gekürzt und das jeweilige Reibe- bzw. Lösungsgeräusch vor dem Vokaleinsatz vollständig abgeklungen. Bei close juncture überlagerte der gelängte Anlautkonsonant teilweise den folgenden Vokal; außerdem war der vorangehende Konsonant gekürzt (z.B. das [ç] in *nicht teuer*). Bei vorangehenden Nasalen wurden Formen mit open juncture in mehr als 80% der Fälle ohne [ʔ] richtig perzipiert. Die Nasale zeigten eine kürzere Dauer und geringere Stimmhaftigkeit als in Anlautposition. Als schwierig erwies sich dagegen die Identifikation der Beispiele, bei denen die Silbengrenze zwischen 2 Vokalen nicht durch [ʔ] signalisiert wurde. Nur 48% der Urteile waren richtig. Der Wechsel der Vokalquali-

³ Die Merkmale decken sich z. T. mit den von Lehiste (3) beobachteten.

tät von [ε] zu [ə], z. B. bei den Formen *Ei-ersatz* und *Eiersatz*, genügte nicht für die Perzeption.⁴

Die Beobachtungen zeigen, daß der Gebrauch von [ʔ] das Erkennen eines Grenzpunktes open juncture (3) sicher gewährleistet. Da aber umgekehrt auch andere Randmerkmale der Vokal- und Konsonantenallophone die Identifikation einer Form open juncture zu einem hohen Prozentsatz ermöglichen, dürfte der [ʔ] für die Perzeption höchstens als redundant bezeichnet werden. Seine Bedeutung erhöht sich, wenn sich Anzahl und Ausgeprägtheit weiterer Randmerkmale verringern. Besonders zwischen vokalischen Lauten trägt er durch seine deutliche Kontrastwirkung zur optimalen Verständigung bei.

LITERATUR

- (1) Korlén, G. und Malmberg, B.: *Tysk Fonetik* (Gleerups, Lund 1960).
- (2) Krech, E.-M.: Sprechwissenschaftlich-phonetische Untersuchungen zum Gebrauch des Glottisschlageinsatzes in der allgemeinen deutschen Hochlautung. *Bibl. Phonetica* No. 4 (Karger, Basel/New York 1968).
- (3) Lehiste, I.: An acoustic-phonetic study of internal open juncture. *Phonetica, Suppl. ad Vol. 5* (Karger, Basel/New York 1960).
- (4) Meyer—Eppler, W.: Grundlagen und Anwendungen der Informationstheorie. *Kommunikation u. Kybernetik in Einzeldarstellungen*, Bd. 1 (Springer, Berlin/Göttingen/Heidelberg 1959).
- (5) Moulton, W. G.: Juncture in modern standard German. *Language* 23: 212—226 (1947).
- (6) Moulton, W. G.: *The sounds of English and German* (Univ. of Chicago Press, 1962).

DISCUSSION

Ezawa:

Es scheint mir hier eine gewisse methodische Tautologie vorzuliegen, wenn man ein Lautmaterial aus dem von vornherein die Varianten landschaftlicher und individueller Art, die zu der gesprochenen hochdeutschen Sprache der Gegenwart unbedingt gehören, ausgeschieden sind, einer Untersuchung zugrunde legt, die die Feststellung der tatsächlichen Sprechwirklichkeit, aus der die adäquaten Aussprachenormen hergeleitet werden sollen, zum Ziel hat. Solche Untersuchung hat freilich ihre praktische Bedeutung, dürfte aber keinen Anspruch darauf haben, auf die gesprochene Sprache als solche Bezug zu nehmen.

MacCarthy:

Question about the identity of the speakers involved in the test, and of the test subjects (listeners), as regards the type of German they normally spoke.

⁴ Für die Herstellung der Sonagramme möchte ich Herrn J. Kyritz vom Institut für Phonetik und Kommunikationswissenschaft der Humboldt-Universität Berlin (Dir. Prof. Dr. G. F. Meier) sehr herzlich danken. Ebenso danke ich Herrn Elektromstr. W. Prescher vom Institut für Sprechkunde der Martin-Luther-Universität Halle-Wittenberg für seine Unterstützung beim technischen Ablauf der Untersuchungen und für die Anfertigung einer großen Zahl von Oszillogrammen.

Ad Eza wa: Es ging in meinen hier dargestellten Untersuchungen nicht darum, den Gebrauch des Glottisschlages in der Sprechwirklichkeit festzustellen, sondern zur Beantwortung der Frage beizutragen, in welchem Maße der coup de glotte im Deutschen zur Erhöhung des kommunikativen Effektes beiträgt. — Handelt es sich andererseits darum, die Aussprache einer bestimmten Sprechstufe zu normieren, dann kann auch nur die Sprechweise solcher Personen den Untersuchungen zugrunde gelegt werden, die diese Sprechstufe tatsächlich realisieren, d. h., Sprecher der Umgangssprache können z. B. nicht maßgeblich für die Sprechwirklichkeit in der Hochsprache sein.

Ad Mac Carthy: Die Sprecher waren Dipl.—Sprecherzieher; sie realisierten die deutsche Hochlautung. Ihre sprachlandschaftliche Herkunft war Norddeutschland. Bei den 57 Perzipienten handelte es sich um Studenten der Fachrichtungen Musikwissenschaft, Germanistik und Sprechwissenschaft. 6 der Hörer waren älter als 30 Jahre, alle anderen waren 19—22 Jahre alt. Von sämtlichen Hörern wurde mit Hilfe gesonderter Fragebogen die sprachlandschaftliche Herkunft ermittelt und mit den Urteilen in Beziehung gesetzt. Es ließen sich jedoch keine Zusammenhänge feststellen, da sich die Fehlinterpretationen etwa gleichmäßig auf sämtliche Hörer verteilten.

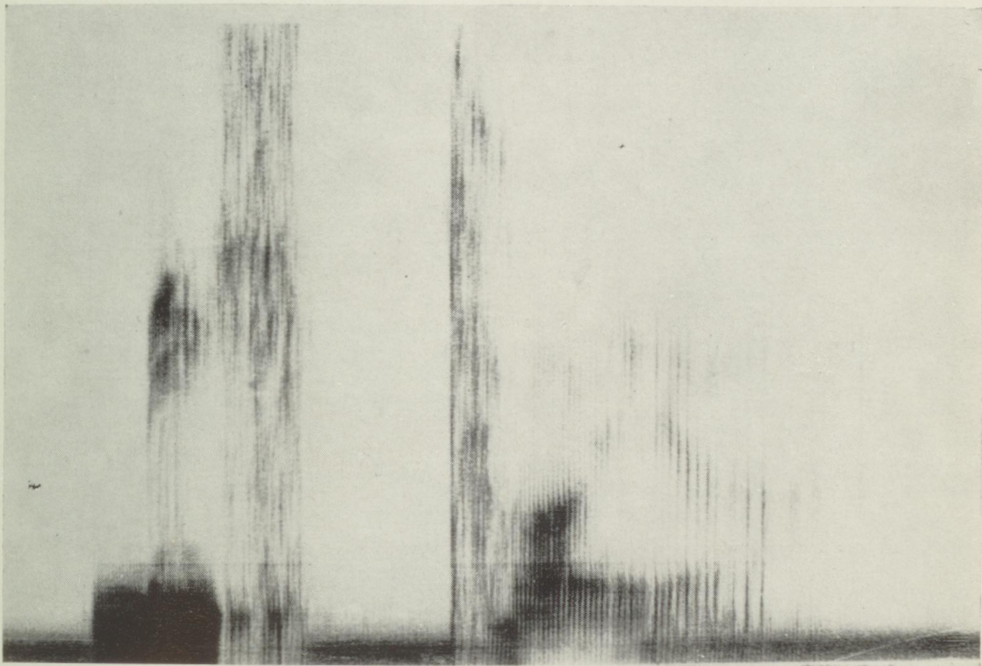


Abb. 1. *nicht teuer*

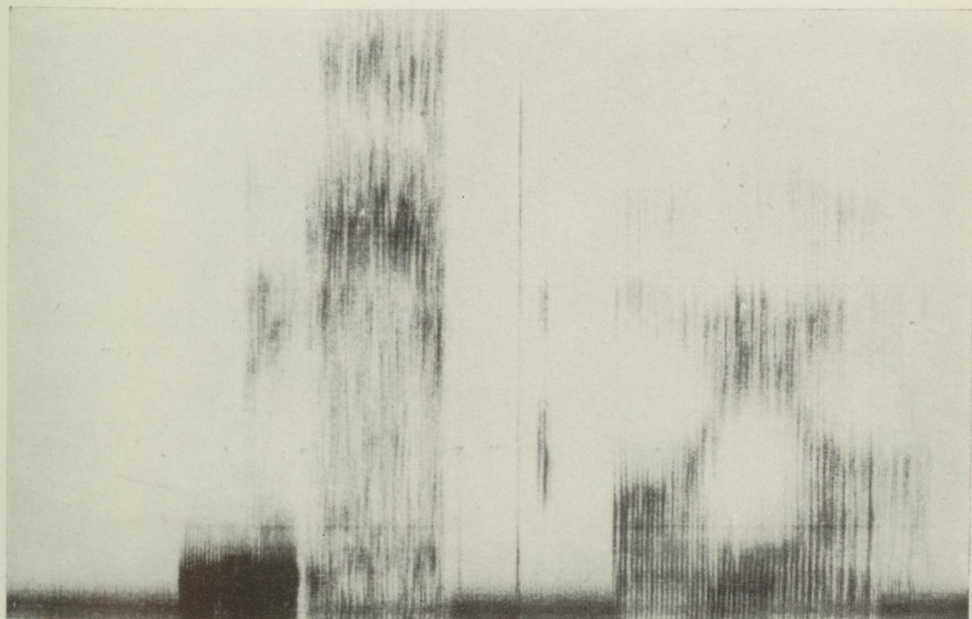


Abb. 2. *nicht euer*

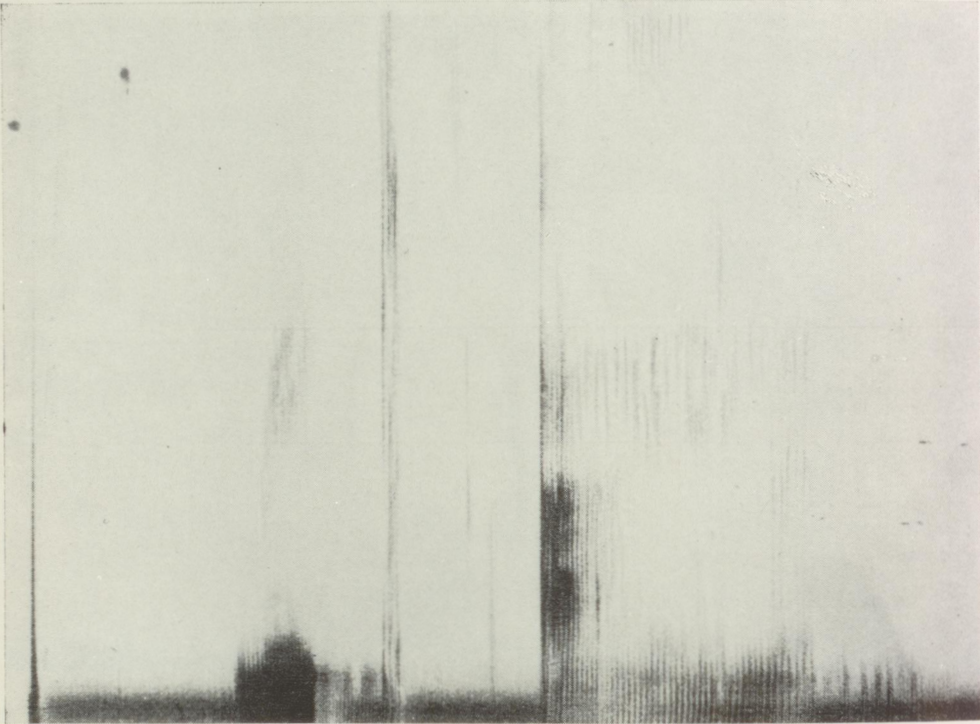


Abb. 3. *mit-eilen*

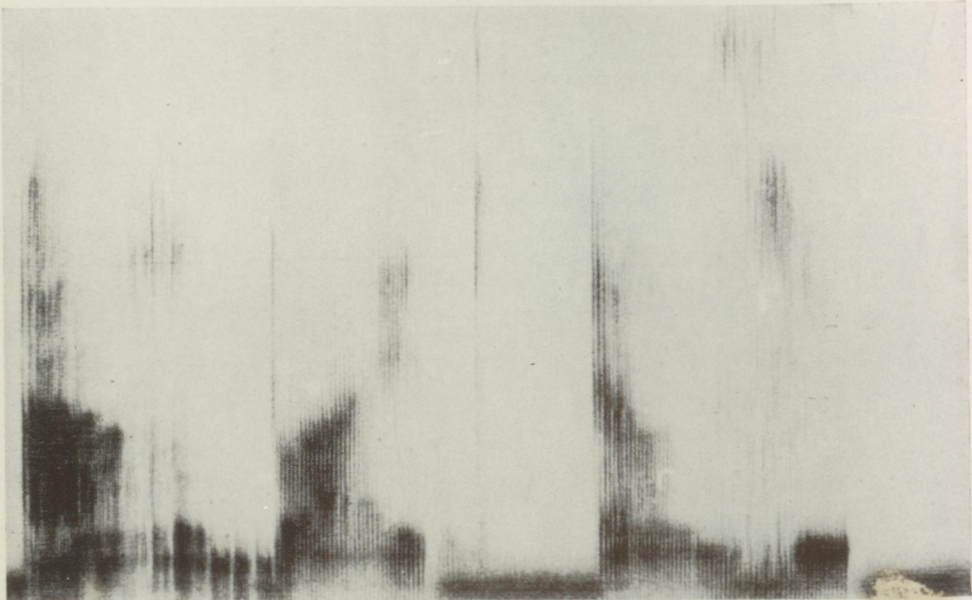


Abb. 4. *Arbeit-ersuchen*

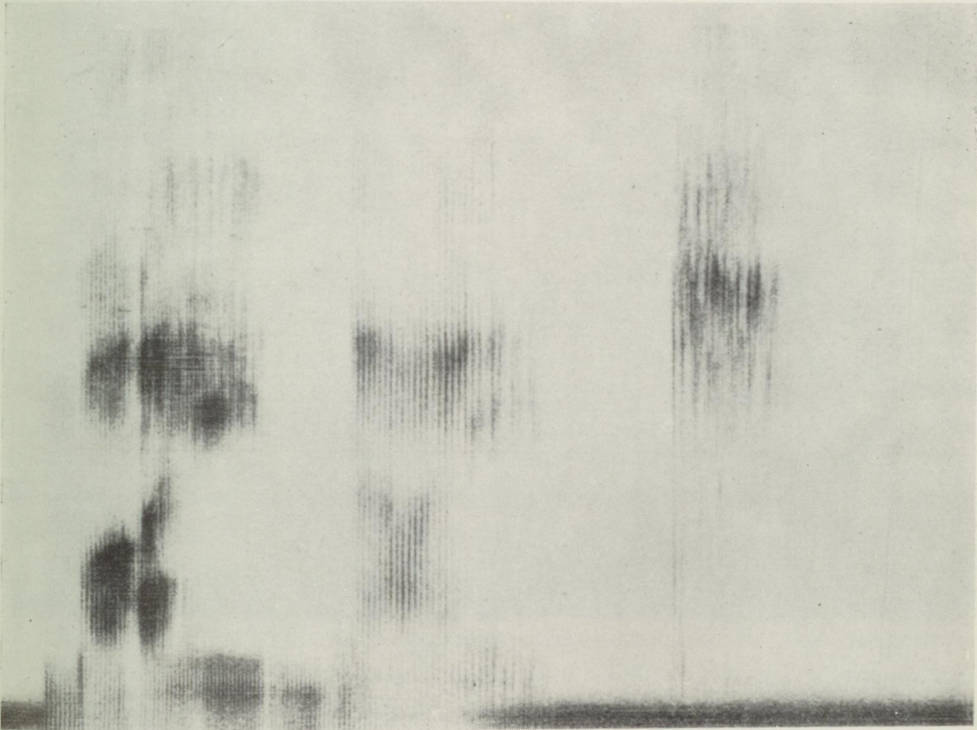


Abb. 5. ¹Meineid



Abb. 6. ¹Mein-eid

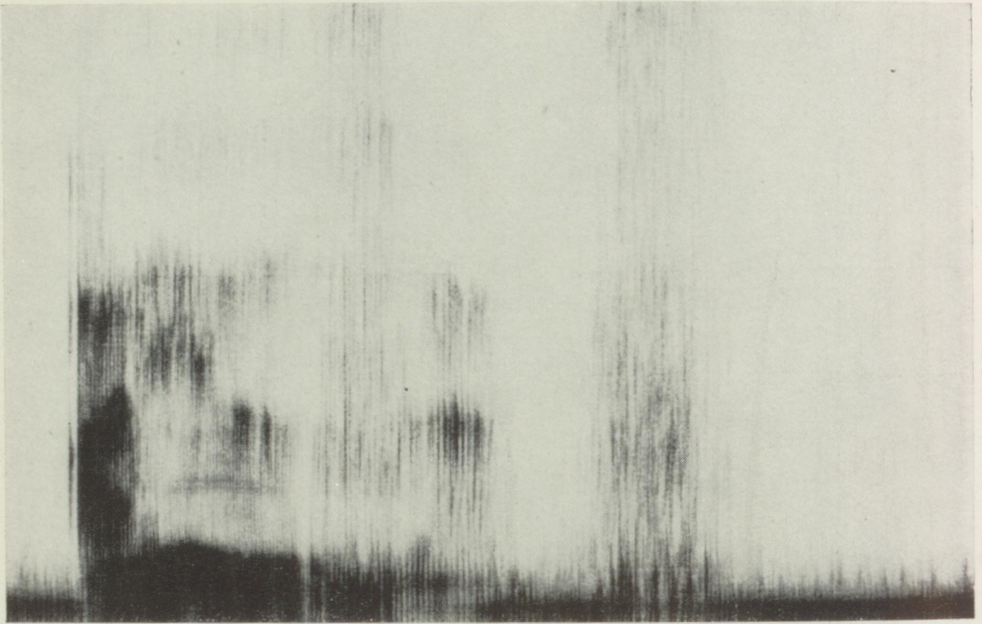


Abb. 7. *ˈEiersatz*

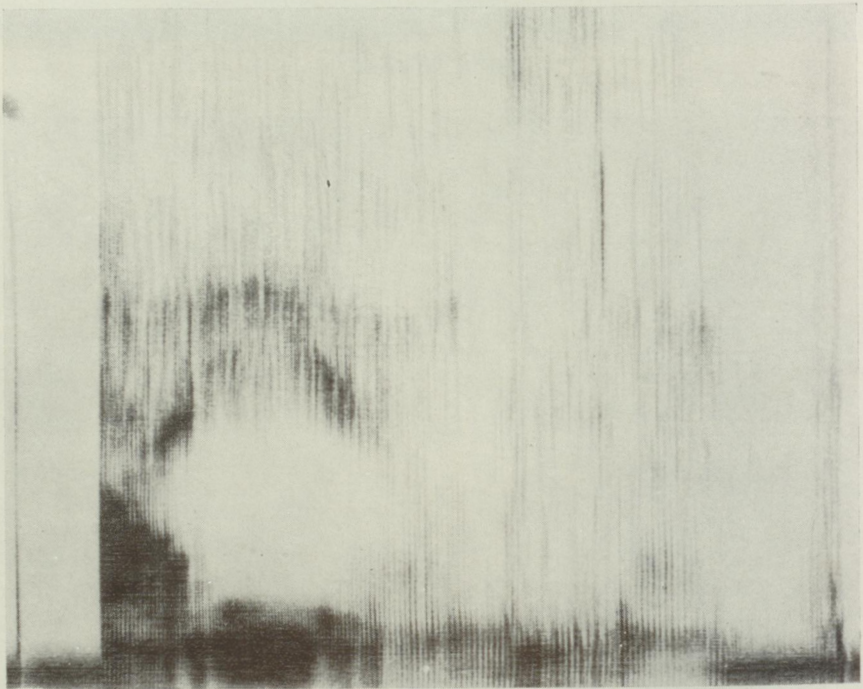


Abb. 8. *ˈEi-ersatz*

ZUR EINPASSUNG ANGLO-AMERIKANISCHER WÖRTER IN DAS ISLÄNDISCHE

BRUNO KRESS

Nach Aufhebung des dänischen Handelsmonopols 1854 entwickelten sich die bis dahin spärlichen Beziehungen Islands zur englischsprachigen Welt. Kaufleute und Hochseefischer lernten Großbritannien kennen, isländische Auswanderer ließen sich in Canada und den USA nieder. Begriffe der Technik, des Sports und des Handels verlangten nach sprachlicher Wiedergabe. Trotz eifriger puristischer Bemühungen nisteten sich eine Reihe englischer Wörter in die isländischen Berufssprachen ein. Die militärische Besetzung Islands im zweiten Weltkrieg durch britische und amerikanische Truppen sowie der Ausbau einer amerikanischen Militärbasis nach dem Krieg führten zu mancherlei Kontakten zwischen den fremden Soldaten und der einheimischen Bevölkerung. Diese Kontakte spiegeln sich in der Entlehnung anglo-amerikanischen Wortguts bestimmter Begriffssphären wider. Das isländische Wortmaterial anglo-amerikanischen Ursprungs ist nicht nur für den Lexikographen und Kulturhistoriker von Interesse, sondern auch für den Phonetiker.

Seit der Besiedlung Islands (870—930) hat die isländische Sprache eine von außen wenig beeinflusste, autochthone Entwicklung durchlaufen. Die geographische Nachbarschaft der Vorstufen des nordgermanischen Isländischen und der Vorstufen des westgermanischen Angelsächsischen war durch die Abwanderung der Angeln, Sachsen und Jüten nach England (5. Jahrhundert) längst beendet. Es ist deshalb auffällig, daß sich der Wiedergabe der einzelnen englischen Phoneme und ihrer Abfolge im Isländischen kaum Schwierigkeiten entgegenstellen. Daß die phonematische Ähnlichkeit des Isländischen mit dem Englischen auf der Grundlage gemeinsamer Herkunft aus dem Germanischen trotz räumlicher Trennung über viele Jahrhunderte erhalten blieb, ist ein Beweis für die relativ große Stabilität von Phonemsystemen.

Das isländische Vokalsystem reicht aus, um alle Vokale des Englischen direkt wiederzugeben oder durch ähnliche zu substituieren: *piece* : *pi:s* [*pi:s*]; *thrill* : *þrill* [*þrɪ:l*]; *to make* : *meika* [*mei:ga*]; *to square* : *skvera* [*sqve:ra*]; *to check* : *tékka* [*tjéhga*]. Kurzes offenes [æ] wird durch das isl. [a] substituiert, das helle Farbe hat: *jack* : *tjakkur* [*tjahgü.r*]. Dafür wird kurzes englisches [a] (mid-back-narrow) durch isl. offenes [ǫ] substituiert: *tough* : *töff* [*töf*]. *car* : *kar* [*ka:r*]; *guy* : *gæi* [*gai:i*]. Für den Diphthong [au] kein Beleg, er wäre im Isl. vorhanden: *á* [*au*]. *shop* : *sjoppa* [*sjöhba*]; *hall* : *hol* [*hø:l*]; für [oi] kein Beleg; *bloke* : *blók* [*blouk*]; *hood* : *húdd* [*hud*]; *spoon* : *spín* (Akk.) [*sbu:n*]. Langes [ɜ:] (low-mixed-narrow), das nur vor *r* der Schrift

vorkommt, wird durch [j + r] ersetzt: girl : *görl* [görl]. Der Endungsvokal [ə] ist im Isl. nicht vorhanden, er wird durch einen der üblichen, in der Flexion verwendeten Endungsvokale [a, i, ü] wiedergegeben, bzw. nach der Schrift mit [ɔ] und [e]: troller : *trollari*; poker : *póker* [pou:ge.r]; motor : *mótor*. Das isländische [i:] (lang und kurz) wird nicht in Anspruch genommen, das Englische hat dieses Phonem nicht.

Bis auf wenige Fälle reicht auch das isländische Konsonantsystem für die Wiedergabe englischer Wörter aus. Den englischen Tenues p, t, k entsprechen, z. T. in Übereinstimmung mit dem Englischen, positionsbedingte Varianten, entweder aspirierte Tenues oder stimmlose unaspirierte Medien. Beispiele für p : piece : *pís* [pi:s]; sport : *sport* [sɔrt]; stupid : *stjúpid* [sdju:bi.d]; shop : *sjoppa* [sjɔhba.]. Beispiele für t : tough : *töff* [töf:]; smart : *smart* [smart]; to strike : *stræka* [sdrai.ga.]. Beispiele für k : (American) *kani* [ka:ni.]; fake : *feik* [fei:k]; to square : *skvera* [sgve.ra.]; to make : *meika* [mei.ga.]; to hook : *húkka* [huhga.].

Den englischen stimmhaften Medien entsprechen stimmlose im Isländischen; nach kurzem Vokal sind sie, der isl. Quantitätsregel entsprechend, gelängt. Beispiele für b : body : *boddi* [bɔd:i.]; snob : *snobb* [snɔb:]. Beispiele für d : to dig : *digga* [dig:a.]; body : *boddi*. Beispiele für g : game : *geim* [gei:m]; to dig : *digga*.

Stimmlose Spiranten (außer [ʃ]) sind im Isländischen vorhanden, auch der behauchte Einsatz von Vokalen. Beispiele: fake : *feik* [fei:k]; to bluff : *blöffa* [blöf.a.]; thrill : *þrill* [þri:l.]; smart : *smart* [smart]; bus : *böss* [bɔs:]; to hook : *hukka* [huhga.]. Der Spirant [ʃ] in shop wird substituiert durch palatalisiertes s + j-Abglitt: *sjoppa* [sjɔhba.]. Schwierigkeiten bestehen bei der Wiedergabe der Affrikata [tʃ], da das Isl. keine Affrikaten kennt. Als Substitutionen dienen [tj], [sj], [s]. Beispiele: to check : *tjekka* [tjehga.]; chance : *sjans* [sjans]; to switch : *svissa* [svi:s.a.].

Die stimmhaften Spiranten [v] und [ð] sind im Isländischen vorhanden, jedoch fehlt [z]. Es wird durch den stimmlosen Spiranten [s] wiedergegeben: civil : *sivil* [si:vi:l]; für [ð] kein Beleg; to squeeze : *skvísá* [sgvi:sa.]. Verständlicherweise bereitet die engl. stimmhafte Affrikata [dʒ] Schwierigkeiten. Sie wird durch [sj], [j], [ds] ersetzt: gentleman : *séntilmáður* [sjendil-]; jeep : *jeppi* [jehbi.]; bridge : *bridds* [brids].

Für den Halbvokal [j] fehlen Belege, er wäre im Isl. vorhanden. Hingegen muß engl. [w] durch isl. [v] ersetzt werden: whist : *vist*.

Die Liquida [l] ist im Isl. nie velarisiert. [r] ist gerolltes Zungenspitzen-r und wird nicht vokalisiert: girl [gə:l] : *görl* [görl].

Die Nasale [m], [n], [ɳ] bieten keine Schwierigkeiten. Sie haben jedoch wie die Liquiden r und l stimmlose, positionsbedingte Allophone (vor p, t, k der Schrift). Das vorgelegte Material weist nur wenige Beispiele auf: gentleman : [sjendil.l.]; smart : *smart* [smart].

Die meisten Entlehnungen fügen sich bereits von Hause aus dem isländischen Akzentgesetz: expiratorisch stärkster und melisch höchster Akzent auf der ersten Silbe. Soweit sich feststellen ließ, werden alle Fremdwörter dem isl. Akzentgesetz unterworfen: sentimental : *séntimental* [sændimeŋda.l].

DISCUSSION

Lieberman:

Borrowings from English into Icelandic (as well as words mispronounced by Icelanders speaking English) can throw additional light on the prosodic system of English. Icelanders pronouncing such words as 'got' are apt to preaspirate the final stop, but nothing of this sort happens in 'god'. The difference between Icel. (ɔ^{ht}) and (ɔ^l) is perhaps that between a word with a 'stød' (preaspiration) and that without it. Could the difference between Engl. 'god' and 'got' not be of similar character?

Blanár:

Die Adaptierung des Lautgebildes des entlehnten Wortes in der annehmenden Sprache ist nur eine Seite des Prozesses der Eingliederung des entlehnten Wortes. Wichtig sind vor allem noch die paradigmatischen und syntagmatischen Beziehungen der neuen Benennungseinheit. Diese Beziehungen ändern sich in der Entwicklung des Wortschatzes. Die graphische Darstellung der relevanten lexikalischen Beziehungen zeigt uns den Platz des entlehnten Wortes im diachronischen Aspekt.

ZUR BEEINFLUSSUNG DER PROSODIE IN DER SPRACHE VON STOTTERERN DURCH SIMULTANE MUSIKPERZEPTION

E. KURKA — E. STOCK

Problemstellung. Die Einwirkung von Musik auf die zwischenhirngesteuerten Funktionen ist exakt nachgewiesen worden (Destunis/ Seebrandt, Stokvis u. a.). Nach Bykow (vgl. Destunis/Seebrandt) wirkt die Musik wahrscheinlich über die Großhirnrinde indirekt auf Thalamus und Hypothalamus, so daß die Rinde auf diese Weise eine regulatorische Funktion auf Herztätigkeit, Atmung und Vasomotorik ausübt. Musik mit raschem Tempo, starken rhythmischen Effekten und entsprechender Dynamik hat meist eine sympathikotrope Wirkung, dynamisch „zurückgenommene“ Musik mit melodisch ausgeprägter Linie sowie klarer, beruhigender Harmonik und Klangfarbe wirkt dagegen im sympathikolytischen Sinne (Destunis/Seebrandt).

Als Ursache für das Stottersyndrom werden in letzter Zeit nicht so sehr pathologisch-anatomische Veränderungen im Thalamus bzw. striopallidären System angenommen, sondern mehr dynamische Abweichungen, die durch starke Affekte hervorgerufen werden und vegetativ gesteuerte, aber kortikal ausgelöste Hypertonien und Hyperkinesen hervorrufen (Seeman). Eine Umstimmung durch geeignete musikalische Einwirkung dürfte daher bei Balbuties-Sprechern als Methode der Wahl in der komplexen Beeinflussung ihren Platz haben. Unsere Untersuchung verfolgte das Ziel, einen meßbaren Nachweis für den empirisch belegten und bereits mehrfach ausgenutzten Effekt zu liefern, daß bei Stotterern, die mit musikalischer Untermalung („musikalischer Kulisse“) sprechen, die Symptome verringert bzw. abgebaut werden (Kurka; Josef; Taubitz; vgl. Wendler-Reuter/Wendler: Sprechstimme beim Melodram).

Untersuchungen. Für die musikalische Untermalung wurden verwandt: 1. der 2. Satz aus Haydns Symphonie Nr. 75 D-Dur und 2. der 1. Satz aus dem Divertimento in A-Dur von Mozart. Beide Musikbeispiele haben einen gelösten Grundcharakter, die Mozartmusik ist jedoch wesentlich schneller, lebhafter, mit stärkerer Dynamik und Rhythmik. Die Probanden lasen einen Versuchstext (7 Einzelsätze und 1 längeren Abschnitt) zunächst hintereinander, und zwar ohne musikalische Untermalung (Lesung O), mit der Haydnmusik (H_1) und mit der Mozartmusik (M_1), sodann eine Woche später wiederum mit der Haydnmusik (H_2) und abermals eine Woche später mit der Mozart-Musik (M_2). Der zeitliche Abstand bei den beiden letzten Lesungen sollte Wiederholungseffekte ausschließen. Die Musik wurde über Kopfhörer auf ein Ohr des Probanden übertragen. Es beteiligten sich 9 Personen (3 weibl., 6 männl.,

Durchschnittsalter 23,6 Jahre). In zwei Fällen konnten je 2 Lesungen nicht durchgeführt werden. Alle Lesungen wurden auf Tonband aufgenommen, aus den Lesungen O, H₁ und M₁ wurde der 3. Einzelsatz mit Tonhöhenschreiber (vgl. Stock) registriert. Es wurde festgestellt (s. Tabelle): 1. die Häufigkeit der Symptome, 2. die Dauer einer Lesung (in sec., abgerundet), 3. das Intervall zwischen höchstem und tiefstem Ton im registrierten Satz (in Vierteltönen, für den Pb. U. Diagramme aus H₂ und M₂), 4. das Mittel der Tonhöhenbewegungen im registrierten Satz (in Vierteltönen über C bei Männern bzw. c bei Frauen).

Tabelle 1

Pb.	Häufigkeit der Symptome					Dauer der Lesung					Intervall zw. Extremwerten			Mittel der Grundtonbewegung		
	O	H ₁	M ₁	H ₂	M ₂	O	H ₁	M ₁	H ₂	M ₂	O	H ₁	M ₁	O	H ₁	M ₁
C.	11	3	4	2	3	90	79	84	89	84	24	22	23	10	10	11
D.	27	8	13	8	15	130	99	92	97	107	14	16	16	19	19	23
E.	3	0	0	—	—	86	85	88	—	—	19	18	18	23	25	26
F.	17	6	9	4	5	121	96	89	86	91	10	14	16	21	19	21
G.	11	12	9	20	21	91	99	95	95	98	13	12	14	16	17	18
H.	43	18	12	12	16	117	104	92	113	99	9	10	12	13	14	14
S.	16	6	9	8	10	108	97	88	92	90	11	14	16	22	23	23
U.	30	—	—	32	14	175	—	—	136	137	6	10	8	27	25	25
V.	69	43	41	14	9	193	157	140	110	93	15	15	12	15	15	16

Ergebnisse. Von allen Probanden wurde übereinstimmend die subjektive Empfindung geäußert, daß der Sprechablauf durch die simultane Musikperzeption wesentlich erleichtert werde. Dabei wurde die Haydn-Musik vorgezogen (von 6 Pb., 3 entschieden sich nicht). Auch objektiv läßt sich der Einfluß der Musik nachweisen. Bei der Bearbeitung der in der Tabelle zusammengefaßten Zahlenwerte mit statistischen Prüfverfahren (Zeichentest, Wilcoxon-test, t-Test für korrelierende Stichproben (Adam) ergibt sich folgendes: Hinsichtlich der Symptommhäufigkeit und der Dauer der Lesung bestehen zwischen den Werten der Lesung O und denen der Lesungen H₁, M₁, H₂, M₂ signifikante Unterschiede (mit 5 %, 1 × mit 10 % Irrtumswahrscheinlichkeit), und zwar wird die Symptommhäufigkeit geringer, ebenso die Dauer als Ausdruck der Symptomschwere. Der bei einigen Probanden sehr deutliche Unterschied im Symptombefall und im Sprechtempo zwischen den Lesungen mit Haydn-Musik einerseits und Mozart-Musik andererseits kann dagegen nicht verallgemeinert werden, hier mußte die Nullhypothese angenommen werden. Die Wirkung der rhythmischen Besonderheiten beider Musik-Sätze wird durch den individuell unterschiedlichen Grad der Empfänglichkeit durchkreuzt. Ein unterschiedlicher Einfluß der beiden Musiken konnte nur in bezug auf das Mittel der Grundtonbewegung statistisch

verifiziert werden. Die Mozart-Musik führt hier zu einer zwar geringfügigen, aber statistisch signifikanten Erhöhung der mittleren Sprechstimmlage, verglichen mit den Lesungen O und H₁ (mit 5 % Irrtumswahrscheinlichkeit). Unterschiede hinsichtlich der Intervallgröße bei den einzelnen Lesungen konnten nicht gesichert werden. So ergibt sich: 1. Während des Lesens perzipierte Musik von der Art der ausgewählten führt bei Stotterern zu einer Verringerung des Symptombefalls und der Symptomschwere. 2. Lebhaft und rhythmisch betonte Musik selbst bei gelöstem, heiterem Grundcharakter hat jedoch affektive und muskuläre Spannungen zur Folge und sollte in Fällen, wo eine möglichst weitgehende Umstimmung zur sympathikolytischen Seite hin erreicht werden soll, nicht verwendet werden.

LITERATURVERZEICHNIS

- Adam, J.: Einführung in die medizinische Statistik (2. Aufl. Volk und Gesundheit, Berlin 1966).
 Destunis, G. und Seebandt, R.: *Beitrag zur Frage der Musikeinwirkung auf die zwischenhirngesteuerten Funktionen des Kindes*. Musik in der Medizin, hg. v. H. R. Teirich, pp. 34—42 (Fischer, Stuttgart 1958).
 Josef, K.: Wirkungen von Musikbetätigung und Musikhören auf Soziabilität und Sprechhemmungen bei Stotterern. *Sprachheilarbeit* 11: 129—136 (1966).
 Kurka, E.: Ein Beitrag zur Übungsbehandlung des Stotterns. *Wiss. Z. Univ. Halle, Ges.-Sprachw.* 11: 1595—1616 (1962).
 Seeman, M.: Sprachstörungen bei Kindern. 2. Aufl., p. 316 (Volk und Gesundheit, Berlin/Jena 1965).
 Stokvis, B.: Psychosomatische Gedanken über Musik. Musik in der Medizin, hg. v. H. R. Teirich, pp. 43—53.
 Stock, E.: Untersuchungen zur Intonation bei Stotterern. *Folia phoniat.* 18: 447—461 (1966).
 Taubitz, R.: Der Tanz in der Stottererbehandlung — seine Heilchancen in psycho-motorischer Hinsicht. *Sprachheilarbeit* 11: 218—223 (1966).
 Trojan, F.: Der Ausdruck der Sprechstimme. 2. Aufl. (Maudrich, Wien/Düsseldorf 1952).
 Wendler—Reuter, U. und Wendler, J.: Zum Verhalten der Sprechstimme im Melodram. *Wiss. Z. Humboldt-Univ. Berlin, Ges.-Sprachw.* 15: 429—438 (1966).

DISCUSSION

Handzel:

Phoniatisches Forschungszentrum in Wrocław führt seit Jahren Untersuchungen zur Etiopathogenese des Stotterns. Die Forschungsergebnisse der behandelten Fälle ergeben zum weitaus größten Teil das der Spasmophilie entsprechende klinische Bild. Die nervöse muskulöse Reizbarkeit ergibt hier Diagramme, die für die Tetanie typisch sind. Diese Forschungen bahnen neue Wege der Behandlung.

Bystrzanowska:

Die Behandlung des Stotterns mit Hilfe der Musik beruht — wie es scheint — auf zwei Faktoren: Vertäubung und Rhythmus. Für die Vertäubung genügt eine Intensität um 10 dB höher als

die Sprechintensität des Stotterers, um ihn zu vertäuben, besonders wenn Lautsprecher benutzt wurden. Es scheint also, daß in diesem Fall der Vertäubungseffekt nicht zu übersehen ist. Zweitens spielt der Rhythmus hier auch eine bedeutende Rolle.

Wendler:

Es wäre interessant zu vergleichen, wie Ihre Patienten auf Sinustöne reagieren. Wir kennen ja alle die Erscheinung, daß Stotterer deutlich weniger Symptome zeigen, wenn sie die Töne eines Audiometers hören. Vielleicht könnte man auf diese Weise den tatsächlichen Einfluß der Musik noch präzisieren.

Kurka:

Die Reaktion der Probanden auf Sinustöne soll als Anregung für weitere Untersuchungen genutzt werden.

Eine Gleichsetzung mit dem Vertäubungseffekt ist deshalb nicht vertretbar, weil die Probanden während der Musikperzeption ihre Sprache hören konnten. Damit ist die Eigensteuerung durch das Gehör gewährleistet, während dies bei der Vertäubung nicht der Fall ist. Wir haben die Übertragung mit einem Kopfhörer durchgeführt, weil wegen der Tonhöhenaufnahmen eine Lautsprecherübertragung in den Raum unterbleiben mußte.

Stock:

Eine Vertäubung lag u. E. nicht vor: 1. weil wir im Eigenversuch die Lautstärke der Musik reguliert haben, 2. weil die Musik nur auf ein Ohr übertragen wurde und 3. weil wir uns während der Übertragung mühelos mit dem Patienten unterhalten konnten. Während des Einhörens in die Musik (bis zu 1½ Min.) beobachteten wir eine deutliche körperliche und mimische Entspannung, so daß wir die Sprachverbesserung auf die emotionalen „Inhalte“ und auf den Musikrhythmus zurückführen können, wie es beim Hören von Sinustönen nicht der Fall ist.

ON THE NECESSITY OF DISTINGUISHING BETWEEN SPEAKING AND LISTENING

HARLAN LANE*

When DeCordemoy wrote in 1668: "Whensoever any sound agitates the Brain, there flow immediately spirits towards the muscles of the Larynx, which duely dispose them to form a sound altogether like that which was just now striking the Brain...",¹ he might have been the first scientist to naively conflate the processes of speaking and listening, but he certainly was not the last. It is evident that speaking and listening are related but distinct processes, and that an adequate theory of language behavior must take these facts into account. Nevertheless, most theories give weight to the presumptive similarities between uttering speech and listening to it, while giving little weight to their differences. Conflationist theories are prominent in such diverse areas as studies of speech communication, at both segmental and suprasegmental levels, automatic speech recognition, foreign language learning, and child language.

With respect to the first domain, Twaddell wrote, in 1952: "...The hearer matches the acoustic stimuli he receives against his own habits of muscular speech action, and identifies the incoming speech sound as corresponding to this or that of his own speech articulations. At both ends of a speech transmission, it is the muscular activity, not the acoustic character, which dominates the identification". Hockett reasoned similarly in his *Manual of Phonology* (1955), Liberman in his review of research on speech perception (1957), Delattre in his survey of the acoustic correlates of consonants and vowels (1958): in short, speech is perceived by reference to articulation.

At the suprasegmental level, Schmitt (1924), Jespersen (1932), D. Jones (1950), Stetson (1951), Ladefoged (1958, 59, 63), Fonagy (1958, 65), Laziczius (1959), Lehiste & Peterson (1959) have all favored a "motor theory" of stress perception, effectively agreeing with S. Jones when he wrote in 1932: „Accent is *sui generis* depending for its perception on the kinesthetic sense... The listener refers what he hears to how he would say it. Thus he translates exteroceptor into proprioceptor sensations, the kinesthetic memory serving as stimulus." Galunov & Chistovich

* Center for Research on Language & Language Behavior, Ann Arbor, Michigan, U.S.A. and Institut de Phonétique de Paris, France.

¹ Because of space limitations, references are available from author.

(1966) and Liberman (1967) express the corresponding position with regard to the perception of pitch and intonation.

Analysis-by-synthesis is a strategy for automatic speech recognition that may simulate the human perceptual process according to Stevens (1960). It is also a stage in second-language learning according to Hockett (1955). Finally, Allport (1924), Liberman et al. (1961), and Kozhevnikov & Chistovich (1965) see it as a component of language development in the child; the latter write: "In the process of imitation [by the child] are provided conditions that are favorable to the forming of conditioned-reflex correlations between groups of sound signals and complexes of articulatory motions... It is assumed that these conditioned reflex correlations play an important part also in the process of speech recognition by an adult."

Opposing the conflation of speaking and listening are, in the first place, the findings of developmental and clinical studies by Lenneberg (1962), MacNeilage, Rootes & Chase (1967), and Fuller (1967); "Neither babbling, imitation nor articulate speech is necessary for the understanding of the natural language", Fuller concluded. Fant (1964), Jakobson & Halle (1956), and Jakobson in his book on child language (1942), all emphasize the relative independence of speaking and listening during language development. A similar view concerning *second* language acquisition is held by Jakobson & Halle (1956), and by Sapon (1965). Morton & Broadbent doubt that the "homunculus is really necessary" in their 1965 paper on passive versus active recognition models, as does Fant (1963, 66).

Further evidence for the functional independence of speaking and listening comes from studies of the perception of the suprasegmental characteristics of speech. Ladefoged (1959) has suggested that we judge loudness in terms of our own vocal level, and Warren (1962) has suggested that we judge our own vocal level in terms of loudness but, in fact, the autophonic scale (the scale of the speaker's perception of his own voice) is not the reception scale, and the reception scale is not the autophonic scale (Lane, 1961, 62, 63).

At the segmental level, the inference that speaking mediates listening is experimentally based on evidence for "categorical perception" of speech sounds. However, noises and visual patterns can be perceived categorically, too (Lane, 1965, 66). In order to illustrate how closely recent findings for color perception (Lane & Kopp, 1967) match those for speech and how, therefore, they constrain the interpretation of the speech results, we may substitute color terms for speech terms in an article on the motor theory of speech perception (Liberman et al., 1963): "Although the [colors] lie on a [visual] continuum, the perception is essentially discontinuous. Because of the discrimination peaks at the [color-class] boundaries, the incoming [colors] are [seen] categorically and they are, therefore, quickly and accurately sorted into the appropriate [color class]. "Although this article's *description* of categorical perception turns out to apply to color as to speech sounds, the same may not be said concerning its *interpretation* of categorical perception: "What kind of mechanism underlies the categorical perception of the [colors]. The answer seems to us... that the perception

of [color] is tightly linked to the feed-back from the speaker's own articulatory movements... In time, these movements... come to mediate between the [color] and its ultimate perception."

DISCUSSION

Singh:

It may be pointed out that not only the consonants are perceived categorically but also the vowels. The results of a recent experiment conducted by Singh and Morehead (reported at the last Acoustical Society meeting in New York) showed sharply dividing spectral patterns in perceiving the consonants (p t k) as well as the vowels (a i u); especially the frequency spectra that facilitated and deteriorated the identification of vowels [i] and [u] were separated, in the lower-end of the spectrum, quite categorically.

DER LEE-EFFEKT VOM GESICHTSPUNKT DER INFORMATIONSTHEORIE

J. LANGOVÁ*—M. MORÁVEK**—A. NOVÁK.***—M. PETŘÍK†

Der Lee-Effekt besteht, wie bekannt, darin, daß die sprechende Person dem Einfluß einer verspäteten Information über die eigene Sprachproduktion ausgesetzt wird. Obwohl die Wirkung des Lee-Effekts auf die Sprache normalsprechender Personen sowie der Stotterer bereits bekannt ist, bleibt noch eine Reihe von Fragen ungeklärt.

An erster Stelle steht die Frage, auf welche Weise das verspätete Signal perzipiert und durch die Zentralmechanismen verarbeitet wird. Die typischste Äußerung einer auftretenden Sprachstörung ist das Wiederholen des bereits ausgesprochenen Lautes, der Silbe oder des kurzen Wortes. Wenn wir uns ein mehrsilbiges Wort in die Elemente ABC aufgeteilt vorstellen, kann man die gestörte Sprache schematisch als ABBC darstellen. Das Element A wird normal ausgesprochen. Die Informationen davon werden durch die Knochenleitung und durch den Rest der Luftleitung zugeführt. Sobald der Sprechende das B ausspricht, hört er die Information über das Aussprechen des B sowie die verspätete Information über das Aussprechen des A. Das Zentrum macht eine Korrektur und es wird von neuem das B ausgesprochen, das nach dem A folgt.

Die Situation ist jedoch komplizierter. Als Beweis dienten uns die Sonagramme, die wir mit Hilfe von 2 Tonbandgeräten hergestellt haben, wobei die Signale gegeneinander verspätet waren wie bei dem normalen Lee-Effekt. Es entstand ein Gemisch von verspätetem und unverspätetem Signal und dieses Gemisch hatte viele Eigenschaften einer gestörten Sprache. Es drückte außerdem verhältnismäßig genau die wirklichen Verhältnisse auf dem Eingang des akustischen Kanals beim Lee-Effekt aus. Der Mensch perzipiert nämlich in dieser Situation das Gemisch beider erwähnten Signale — des verspäteten sowie des unverspäteten — und keinesfalls dieses oder jenes Signal isoliert. Deshalb müssen wir die Frage stellen, ob sich das Zentrum der wahrgenommenen Information nicht nur einfach anpaßt und ob es nicht die ganze sprachliche Äußerung nach den Informationen korrigiert, die dieses Gemisch bilden.

* Foniatriká klinika FVL KU Praha, přednosta prof. MUDr. M. Seeman, DrSc.

** Ústav leteckého zdravotnictví, Praha.

*** ORL katedra IDVLF, Praha.

† Foniatriká laboratoř FVL KU Praha.

Es ist nämlich nicht ausgeschlossen, daß hier zum Teil die sogenannte Erscheinung von Tomatis zur Geltung kommt. Diese Erscheinung beruht darauf, daß der Mensch den Charakter seiner stimmlichen Äußerung ändert, jenachdem ob er sich normalerweise hört oder ob seine Stimme durch eine Reihe von Filtern passiert und — um einige Frequenzbänder reduziert — den Eingang des akustischen Kanals erreicht. In diesem Fall unterdrückt der Sprechende in seiner Stimme die gleichen Frequenzen, die durch die Filter unterdrückt wurde. Mit anderen Worten gesagt: Die sprachliche Äußerung wird dem Charakter der Informationen angepaßt, die durch die akustische Rückkoppelung zugeleitet werden.

Die weiteren Ergebnisse unserer Versuche mit dem Lee-Effekt zeugen davon, daß unter normalen Verhältnissen das Sprachsignal als eine komplizierte Struktur wahrgenommen wird. Wir haben nachgewiesen, daß bei Personen, die auf den Lee-Effekt empfindlich reagieren, entweder Sprachstörungen auftreten, oder wenn es sich um Stotterer handelt, eine Verbesserung der Sprache zutage tritt, falls durch Filtration oder durch ein starkes Geräusch der Teil der Information, der als semantisch bezeichnet werden kann, praktisch vollkommen unterdrückt wird. In einem solchen Fall greift in die sprachliche Äußerung nur die rhythmische und die melodische Komponente ein, die bei Geräuschanwendung auf den Sonagrammen fast nicht zu unterscheiden ist. Ihr Vorhandensein im akustischen Kanal genügt, um bei den Versuchspersonen einen signifikanten Abfall des Sprechtempos hervorzurufen. Dies gilt als Nachweis, daß nicht die Nebenwirkungen (wie Filtration oder Geräusch) wirken, sondern die Verspätung der stark reduzierten akustischen Information. Die Vorgänge, die am Verlauf der Sprache in der angeführten Situation teilhaben, werden auf relativ niedrigem subkortikalem Niveau realisiert, höchstwahrscheinlich im striopallidären und thalamischen System.

Die Sprache der Stotterer kann, zumindest im Zusammenhang mit unserer Thematik, als experimentelle Situation aufgefaßt werden. Dabei übernimmt das durch die akustische Rückkoppelung mit Verspätung perzipierte Sprachsignal die Rolle eines korrigierenden Faktors, unter dessen Wirkung die gestörte Sprache fast normalisiert wird. Es ist sehr schwierig, diese Tatsache zu erklären. Bei den Stotterern ist die Art der Perzeption des Sprachsignals und seine Verarbeitung wahrscheinlich anders als bei normalsprechenden Menschen. Einige Phänomene zeugen davon, daß das akustische Signal nur als Bestandteil des akustischen Mediums perzipiert wird und vor allem in der Zeitspanne des sogenannten Präphonations- oder Initialtonus als akustische Prothese wirkt, die dem Zentrum zur Realisation der Phonationsanstrengung genügt.

Es kann auch nicht ausgeschlossen werden, daß der Lee-Effekt vor allem deshalb bei den Stotterern in Wirkung tritt, weil aus dem Sprachsignal die musikalische Komponente perzipiert wird und das Zentrum auf Grund dieser Information die rhythmische und die melodische Seite der Sprachäußerung korrigiert, die bei den Stotterern wesentlich gestört ist.

Aus den oben angeführten Erfahrungen über den Lee-Effekt geht hervor, daß

das Sprachsignal als komplizierte Ganzheit wahrgenommen wird, aus der die Zentralmechanismen in verschiedenen Situationen unterschiedliche Komponente absondern und diesen verschiedene Wichtigkeit beimessen. Die Rückkoppelungsmechanismen, die das neurophysiologische Substrat der Perzeption und der Verarbeitung des Sprachsignals bilden, sind in ihrer Komplexität sowie in ihrer Stellung in der Hierarchie der Nervenzentren nicht einheitlich. Unserer Meinung nach ist es daher notwendig, zu den Funktionsstrukturen des zentralen Nervensystems auf den verschiedenen Ebenen, vom rubronigralen System bis zu den Zonen der Hirnrinde, verschiedene Komponenten des Sprachsignals, vom Intensitätsfaktor bis zum semantischen Inhalt hinzuzufügen.

THE SYNTHESIS OF COMPONENTS IN VOICE QUALITY

JOHN LAVER*

It has been said that the distinction between phonetic quality and personal quality, or voice quality in my usage of the term, is one of the basic assumptions of phonetics.¹

Although we take phonetic quality as the basic datum of our subject, we know less than we might about voice quality, and I have been trying to construct a simple descriptive model of this area.

The essence of the approach is that voice quality is analysable into components, in much the same way that a consonant is described by reference to the voicing, place and manner components in its production.

Briefly, voice quality is conceived as deriving, acoustically, from two main sources: firstly, the permanent and ephemeral organic foundation of the speaker's anatomy and physiology; and secondly, the long-term muscular adjustments, or "settings",² once voluntarily acquired perhaps, now unconscious, of the larynx and supralaryngeal vocal tract.

This second section, the long-term muscular settings of the larynx and supralaryngeal vocal tract, is an area open to traditional phonetic techniques of imitation and kinesthetic introspection, and one can thereby form hypotheses about the physiological actions necessary to produce particular qualities heard from a speaker. These hypotheses can be tested by acoustic synthesis, and I have carried out some preliminary simulation of some voice qualities, using PAT, the Edinburgh University resonance analogue speech synthesiser designed by Walter Lawrence.³

Different voice qualities were synthesised as backgrounds to a standard linguistic message, a longish sentence of synthetic speech from "The North Wind and the

* Department of Phonetics, University of Edinburgh.

¹ Ladefoged, P.: "The Perception of Vowel Quality" Ph. D. Thesis, University of Edinburgh, 1960.

² Honikman, B.: "Articulatory Settings" in "In Honour of Daniel Jones", D. Abercrombie et al. (eds.), Longmans, London, 1964.

³ Anthony, J. and Lawrence, W.: "A Resonance Analogue Speech Synthesiser", *Proceedings of the 4th International Congress on Acoustics*, Copenhagen, 1962.

Sun",⁴—"Then the North Wind blew as hard as he could, but the more he blew, the more closely did the traveller fold his cloak around him, and at last, the North Wind gave up the attempt."

Three different fundamental frequency ranges were used, which I shall call Voice, (50—250 c/s), Falsetto, (120—480 c/s), and Creak, (32—128 c/s).

Besides the normal unmodified larynx and vocal tract settings for PAT, some extra optional laryngeal and supralaryngeal modifications were superimposed on the three fundamental ranges. Laryngeal modifications were: harshness and whisperiness. Supralaryngeal modifications included nasalisation; a latitudinal distortion of the vocal tract,—velarisation; and a longitudinal distortion of the vocal tract,—raised larynx.

Thus voices with up to five different components can be simulated: optional harshness, optional whisperiness, and optional nasalisation; one of the three configurations of the vocal tract,—unmodified (i.e. normal), velarised, or raised larynx; and one of the three fundamental frequency ranges, voice, falsetto or creak.

All possible combinations of these components give a total of seventy two different voice qualities, such as: Nasalised Falsetto, Harsh Whispery Creak, Harsh Nasalised Velarised Voice, Whispery Nasalised Raised Larynx Falsetto, and so on. I'd like to play you some of these, reserving acoustic details of the syntheses for later, should anyone be interested. I would especially appreciate any comments on the naturalness or otherwise of the qualities,—judging naturalness by a readiness, for example, to ascribe characteristics of personality to the voice in question.

List of items on Demonstration Tape

1. Normal Voice
2. Normal Falsetto
3. Normal Creak
4. Normal Whispery Voice
5. Normal Harsh Voice
6. Normal Harsh Whispery Voice
7. Normal Nasalised Voice
8. Velarised Voice
9. Raised Larynx Voice
10. Whispery Velarised Creak
11. Harsh Whispery Nasalised Velarised Falsetto

In conclusion, I'd like to say that a componential description seems to be a useful approach to structuring this area. In the synthetic work so far, only categorical

⁴ Uldall, E.: "The Synthesis of a Long Piece of Connected Speech", Stockholm Speech Communication Seminar, 1962.

differences between voices have been attempted; the synthesis of scalar, quantitative differences within a component is the eventual aim of the project.

DISCUSSION

Fant:

1. Have you made an acoustic analysis of various pathological voices?
2. Your nasal quality might have benefitted from a simultaneous shift up in F_1 .

Mazzarella:

The artificial voices used in the demonstration often seemed to suggest a certain age of the speaker.

Laver:

ad Fant: I thank Dr. Fant for his suggestion about raising F_1 to give a better nasal quality.

I have been less concerned with pathological voices than with the voice quality variations which are susceptible of imitation by normally equipped speakers — i.e. voluntary long-term muscular settings of the larynx and vocal tract. I have personally made no acoustic analyses yet, relying more on the somewhat restricted literature, and on physiologically-based hypotheses.

Clearly the next step is to proceed to the acoustic analysis of individual voices.

CERTAIN PHONETIC TENDENCIES PERCEIVED IN THE IDIOLECTS OF SELECTED NATIVE TEXANS

TELETÉ ZORAYDA LAWRENCE*

The study of distinctive phonetic characteristics in the speech of Texas has been, and continues to be, a subject of investigation. A primary study analyzed the responses of 20 Texas-born and educated university student-informants (Lawrence, 1963).¹ Identical material, 160 sentences exemplifying the vowel forms of American English, was used to determine the phonic forms and features inherent within the various idiolects at the time of interview.² Subsequent observation and analyses substantiate and support the initial investigation.

Scholars have observed that in English the vowels provide the "main vehicle for dialectal differences" (Fry, 1964; Kurath and McDavid, Jr., 1961). The speech of Texas, a deservedly unique dialect of American English geographically situated in a transition area, is accordant with this observation. Numerous phonetic changes are encountered on vocalic forms and certain phonetic tendencies are operant (Malmberg, 1963).

The front vowels [ɪ, ε, æ] demonstrate instability within certain phonic contexts. Here a tendency is evidenced toward an interchange of these particular forms with, however, no apparent confusion in word meaning. The most persistent phonic interchange is raising [ε] > [ɪ] before nasals /n, m/ and consonantal /s, t, d, k/. Conversely, [ε] may be lowered to [æ] within similar sound contexts. Less regularly [ɪ] may be lowered to [ε], usually preceding /n, m/. The most unique variation is the lowering of [ɪ] > [æ] before the velar-nasal [ŋ]. This change does not appear to be a phonetic accident. It was noted in 45 % of the cases in the primary study and continues to occur with regularity in subsequent analyses. Due to the inconsistency of sound shifts, frequently unpredictable even within one idiolect, the chief phonetic characteristic is the *interchange itself*.

The tendency toward raising and tensing certain vocalic forms is routinely percei-

* Fort Worth, Texas, Texas Christian University.

¹ Primary study supported by Faculty Research Grant, Texas Christian University, Fort Worth, Texas, 1961—1963. Subsequent observations and analyses conducted at the same institution.

² Vocalic forms [ɪ ɪ ε æ u ʊ ɔ ə ə ~ Δ eɪ oʊ aʊ aɪ oɪ ju] in Grant Fairbanks, *Voice and Articulation Drillbook* (2d ed. rev.), pp. 28—56. Harper and Bros., New York. 1960.

ved. This is more apparent on monophthongal /ɪ, ε, æ/, the first element of diphthongal /au, ai/ and the last element of complex /ju/. A multiplicity of phone types are inherent *within* each particularized symbolic form as exemplified:

- [ɪ] > [ɪ¹, ɪ², ɪ³; ɪ¹:, ɪ²:, ɪ³:]; [ε] > [ε¹, ε², ε³; ε¹:, ε²:, ε³:];
 [æ] > [æ¹, æ², æ³; æ¹:, æ²:, æ³:];
 [au] > [æu, æ¹u, æ²u, æ³u; æ¹:u, æ²:u, æ³:u];
 [ai] > [a¹, a²; a¹:, a²:];
 [ju] > [ju¹, ju², ju³; ju¹:, ju²:, ju³:].

The acoustic factor of half-long to long prolongation is noted. However, the variable of noticeable nasality (nasalization) — “at least qualities which are interpreted as being nasal”—is the most characteristic accompanying feature of this specific tendency (for quote, Van den Berg, 1962). The nasalized acoustic impression is undoubtedly due to the increased localized muscular tension evidenced physiologically in production and not of necessity the acoustic result solely of an insufficient closure of the velopharyngeal sphincter. To illustrate: in the /ɪ, ε/ variants, the blade of the tongue is tensed and the oral resonating cavity is constricted; in the /ju/ variants, hypertensive palatalization occurs; in the [æ, au] variant forms, the dorsum of the tongue is unduly raised, tensed and usually retracted with an accompanying factor of hypertensive musculature in the velar, faucal and pharyngeal areas. All of these factors contribute adversely to the realization of the optimum resonance potential. This suprahyoid hypertension has a pervasive tensing effect upon the subhyoid musculature and, it is hypothesized, upon the *internal* as well as the external laryngeal musculature; thus, it becomes a prime etiological element in the perceived “nasality” (Greene, 1964). With this hypertensive syndrome operant, the complex overtone spectrum has been acoustically altered by certain formant shifts on these particular phonemes (Luchsinger and Arnold, 1965). The term “phonemic nasality (nasalization)” has been suggested to describe this accompanying feature of functional etiology perceived on certain raised and tensed vocalic forms (Lawrence, 1967).

Conversely, a tendency toward backing and/or lowering is noted on vocalic forms /ɔ, v, a/, especially before /r/ in the centering diphthongal combinations [ɔr, vr, ar] > [ɔ⁻r, v⁻r, a⁻r]. Again, a hypertensive syndrome is encountered in the excessive retraction of the tongue and undue muscular tension in adjacent areas, notably the pharynx. The optimal resonance capabilities, both supra- and subhyoid, are not achieved. In fact, the entire muscular synergy is disrupted with its accompanying adverse acoustic effects and the ultimate phonation goal of an optimum pneumo-phonetic balance is not attained.

Also relative to vocalic forms /ɔ, v, a/, a pronounced tendency toward an interchanged usage is evident without limitation to surrounding sound contexts. Again, no confusion in word meaning is apparent.

A marked tendency toward the use of more complex forms is perceived wherein a monophthong is diphthongized or triphthongized. In addition, a diphthongal form

may become a triphthong or a more complex 4-element vocalic form. This phonetic change occurs frequently on /ɪ, ε/, æ before nasals /n, m/ and before consonants /s, t, d, l/. It also is commonly encountered on diphthongal forms /eɪ, oɪ, aɪ/. Exemplified forms are:

- [ɪ] > [ɪ², ɪ³, ɪ⁴; ɪ²:, ɪ³:, ɪ⁴:];
 [ε] > [ε², ε³, ε⁴; ε²:, ε³:, ε⁴:];
 [æ] > [æ², æ³, æ⁴; æ²:, æ³:, æ⁴:];
 [eɪ] > [a²eɪ, a³eɪ, e²eɪ, e³eɪ, e⁴eɪ, a²eɪ, a³eɪ, a⁴eɪ];
 [oɪ] > [e²oɪ, e³oɪ, a²oɪ, a³oɪ, o²oɪ, o³oɪ];
 [aɪ] > [æ²aɪ, a³aɪ, a⁴aɪ, a²uɪ, a³uɪ, a⁴uɪ].

The relationship to the acoustic variable of duration is apparent. However, due to the changes in pitch and timbre perceived during the emission of these complex forms, the effect of this tendency upon the overall rhythm and melody pattern is of even greater import.

Again, conversely, a tendency toward the monophthongization of a more complex form is demonstrated on phonemic /ai, oi/ as the initial element is retained with half-long to long prolongation and the second element is omitted. The lax /ai/ form is rarely perceived; rather, [ar] > [a¹, a²; a¹:, a²:] with additional accompanying features. The phonic form [ɔr] > [ɔ¹, ɔ²; ɔ¹:, ɔ²:] preceding /l, s, z/ with backing on the first prolonged element frequently perceived.

The speech of Texas is an area for research which warrants investigation in greater depth and scope. Certain acoustical and physiological findings already apparent have far-reaching implications for clinical phoniatric procedures.

REFERENCES

- Fry, F. D. (1964), “Experimental Evidence for the Phoneme,” in *In Memory of Daniel Jones*. D. Abercrombie, D. B. Fry, P. A. D. MacCarthy, N. C. Scott, and J.L.M. Trim (eds.), pp. 67—68 Longmans, London.
 Greene, M. C. L. (1964), *The Voice and its Disorders* (2nd ed.), pp. 183—185. J. B. Lippincott, Philadelphia.
 Kurath, H. and McDavid, Jr., R. I. (1961), *The Pronunciation of English in the Atlantic States*, p. 2. The University of Michigan Press, Ann Arbor, Michigan.
 Lawrence, T. Z. (1963), “An Analysis of the Speech of Twenty Students at Texas Christian University.” Unpublished Master’s thesis, Texas Christian University; off-set lithography.
 Lawrence, T. Z. (1967), “Regional Speech of Texas: A Description of Certain Paralinguistic Features.” Proceedings (to be published in 1968, by the Publishing House of the Academy of the S. R. of Romania), Xth International Congress of Linguists, Bucharest, Romania.
 Luchsinger, R. and Arnold, G. E. (1965), *Voice — Speech — Language Clinical Communicology: Its Physiology and Pathology*, p. 458. Wadsworth Pub. Co., Belmont, California.
 Malmberg, B. (1963), *Phonetics*, pp. 101—103. Dover Publications, Inc., New York.
 Van den Berg, J. W. (1962), “Modern Research in Experimental Phoniatrics,” Main Lecture. XIIth International Congress of Logopedics and Phoniatrics (Padua), *Folia Phoniatrica* 14: 117—118.

DISCUSSION

Annan:

The allophonic variations within vowels in given contexts will, I consider, be similar and will not basically affect the system, e.g. [I + ŋ] > [æ̃] ∴ [æ + ŋ] > [æ̃ɪ].

McDavid:

It would be desirable to know the homes of the selected informants and the distribution of these phonetic features according to regions of Texas. Even if the conclusion is reached that these features occur in all parts of Texas, we should first have the data by informants and communities before we generalize.

MacCarthy:

You stated, regarding the tendency to nasalise vowels, that often the soft palate was not lowered. Were the contexts of the vowels in question pre-nasal consonant (in which case presumably the soft palate *would* have been lowered), and if not, what is your evidence for your statement?

Lawrence:

ad Annan: The allophonic variations within vowels in given contexts will be similar and therefore will not basically affect the system; thus, your deduction could be expected to be valid. However, it must be pointed out that the tendency toward *raising* is routinely associated with the acoustic impression of nasalization and not the lowering signal as indicated in your example.

ad MacCarthy: The acoustic impression "interpreted as being nasal" (quote from Van den Berg) is routinely perceived on vocalic forms *not* in pre-nasal position; thus, the familiar phenomenon of assimilative nasality would not be suspect as the causative factor. My evidence for the described hypertensive syndrome being the primary etiological factor involved in this specific tendency, that it is not of necessity the acoustic result *solely* of an insufficient closure of the velopharyngeal sphincter, is derived primarily from practical clinical experience. However, conclusions advanced by Van den Berg and Greene (see above *References*) are highly supportive.

ad McDavid Jr.: The data you mention exists in the body of the primary study relative to the 20 student-informants comprising that particular study. The same described phonetic tendencies continue to be demonstrated by subsequent student-informants born and educated in numerous additional areas of Texas. The usage-patterns are on a highly individualized basis, not all informants employing all tendencies. In addition, specific tendencies do not appear to be confined to specific regions. These observations are admittedly but a beginning and serve as an indication of the magnitude of the task ahead before definitive conclusions can finally be drawn for this vast geographical area.

CLINICAL EVIDENCE AGAINST THE MOTOR THEORY OF SPEECH PERCEPTION

YVAN LEBRUN*

In recent years a number of linguists, including A. Liberman (1957) and P. Denes (1965), have put forward a theory to the effect that "speech is perceived by reference to articulation, that is, the articulatory movements and their sensory effects mediate between the acoustic stimulus and the event we call perception"¹.

The basic idea of this motor theory of speech perception is not new. As far back as 1925, L. Stein wrote: "Wir können ja oft nur dann einen Laut richtig hören, wenn wir auch die kinästhetische Vorstellung desselben haben und in die Tat umsetzen können." And in 1932 S. Jones suggested that "the listener refers what he hears to how he would say it. Thus he translates exteroceptor into proprioceptor sensations, the kinaesthetic memory serving as stimulus."

There are, however, a number of facts which do not fit in well with the motor theory of speech perception. Let us examine some of them:

A few years ago, F. Lhermitte and coll. (1963) reported the case of a 43-year-old woman who could hardly move and could no longer speak though she had no paralysis. „Or, the authors write, il nous semblait que la malade avait peut-être une activité psychologique bien supérieure à celle que son état moteur laissait paraître et qui avait l'apparence d'un coma ou d'une profonde léthargie. Et, comme elle était capable de fléchir l'avant-bras sur le bras, nous avons utilisé ce mouvement comme étant le moyen pour elle d'exprimer une réponse affirmative. Nous lui avons donc dit que nous lui poserions différentes questions et nous lui avons demandé de fléchir l'avant-bras lorsqu'elle voudrait répondre par l'affirmative à une question. Ce code nous a permis de découvrir des capacités intellectuelles dont on ne pouvait prévoir la persistance. Lui ayant demandé, par exemple, de fléchir l'avant-bras lorsque son prénom serait prononcé, une liste de prénoms a été énoncée, un silence de deux secondes environ séparant l'émission de chaque prénom pour laisser à la malade le temps d'effectuer le geste-réponse. La malade a parfaitement reconnu son prénom „Jeanne“. Il en a été de même de son nom de famille. Et ceci, malgré l'énoncé de prénoms et de noms

* From the Neurolinguistic Research Unit, Neurosurgical Department, University of Brussels (Prof. Dr. J. Brihaye).

¹ A. Liberman (1957).

phonétiquement voisins des siens (par exemple Prévost, Prouvot, Pourot, alors que son nom était Pouvreau). Elle a reconnu le nom de l'actuel Président de la République et celui de son prédécesseur, dans une liste de noms d'hommes célèbres. Elle a correctement répondu par le même geste de flexion de l'avant-bras lorsqu'on lui a demandé d'indiquer la date de la „Libération de Paris“ (1944) dans une série de nombres à 4 chiffres. Elle a été capable d'effectuer les opérations simples telles que $4 + 9$, $5 + 8$, 7×5 , etc., en fléchissant l'avant-bras lorsque l'observateur énonçait le chiffre correspondant au résultat correct de l'opération. Il en a été de même pour le problème suivant: j'achète un paquet de cigarettes qui coûte 150 francs, je paie avec un billet de 500 francs; combien me rend-on?“

Some years previously, H. Fang and J. Palmer (1956) had reported a somewhat similar case: “So-called ‘akinetic mutism’... was observed in case 1 of our series and lasted for approximately four days. During that period the patient was not comatose, but he showed flaccid quadriplegia and facial diplegia. He was totally anarthric and unable to move his head or neck. He was able to respond to simple verbal commands by opening and closing his eyes; the frequency of eye-blinking and the pauses between blinks were coded so as to convey slightly more complicated commands. Thus, the only remaining avenue of communication between this patient and his outside world was his ability to close his eyelids on volition.“

In the case reported by Lhermitte and coll. the akinetic mutism was due to a softening of the reticular formation. In the case reported by Fang and Palmer, it was also due to a brainstem disease.

Now, if “the articulatory movements and their sensory effects mediate between the acoustic stimulus and the event we call perception”, how did these speechless patients manage to understand what they were told?

It might of course be argued* that in the adult speaker “the reference to articulatory movements and their sensory consequences must somehow occur in the brain without getting out at the periphery”.² Accordingly, the two patients described above could understand what they were told because their brainstem lesion did not prevent them from translating the acoustic stimuli into their motor equivalents at a higher level, i.e. in the cortex.

The following case history will show up the weakness of this argument:

On September 11th, 1965, late in the evening, Joseph V. was assailed by bandits who severely beat him. Joseph succeeded in reaching a hotel whose owner called for a doctor. The physician arrived and found a stupefied patient with several head wounds. The doctor had Joseph taken to hospital. The next morning, they discovered that Joseph, though he was quite in his senses, could not speak a word. No other neurological symptom could be discovered, however. The patient was referred to

another clinic for further investigation. Joseph wrote his daughter that he was lying in hospital because he had had a fall and he told her not to worry. A week after the aggression, the patient's health condition suddenly deteriorated: Joseph was in a fever and he complained in writing about violent headaches. A röntgenographic investigation of the skull revealed a temporo-parietal fracture. On September 23rd, Joseph was trepanned: his wound was cleansed and the bone fragments tied together. On September 27th, a neurolinguistic investigation revealed that the patient's inner speech and gestural language were intact. The patient could read and write correctly. He had no oro-pharyngeal apraxia and no paralysis. He understood everything one told him but could not utter a single word: all his attempts at speaking resulted in unintelligible grunts. In short, Joseph had a cortical anarthria without any concomitant dysphasic impairment; his mutism was solely due to a general disturbance of his motor speech patterns; Joseph could not speak because the cortical programs that cause words to be vocalized were completely disordered. As I have shown elsewhere³, this diagnosis has been corroborated by the results of follow-up examinations of the patient.

The three cases adduced above indicate that as a result of a brain injury one may lose one's ability to utter words without at the same time losing one's ability to understand speech. This dissociation strongly suggests that motor speech patterns are by no means a necessary link “between the acoustic stimulus and the event we call perception”.

Yet another case which points to the same conclusion has been reported by E. Lenneberg. This author had the opportunity to examine a young boy who was able to answer syntactically complex questions and to obey even tape-recorded verbal orders though he had never been able to speak. This well-documented case shows that one may learn to understand speech even though one is unable to develop vocal skills.

It follows from the above that the motor theory of speech perception, if it be any good at all, “accounts for only one of several ways in which (speech) perception is established”, as P. Denes himself admitted at the preceding congress.

REFERENCES

- P. Denes (1965): On the Motor Theory of Speech Perception. *Proceedings of the Fifth International Congress of Phonetic Sciences* Basel, Karger: 252—258.
H. Fang and J. Palmer (1956): Vascular Phenomena Involving Brainstem Structures. *Neurology* 6: 402—419.
S. Jones (1932): The Accent in French — What is Accent? *Le Maître Phonétique* 40: 74—75.
Y. Lebrun (1967): Remarques sur la dysarthrie. *Acta Neurologica et Psychiatrica Belgica* 67: 835—839.

³ Lebrun (1967).

* See Fry's remarks in the discussion below.

² A. Liberman (1957).

- E. Lenneberg (1962): Understanding Language Without Ability to Speak: A Case Report. *Journal of Abnormal and Social Psychology* 65: 419—425.
- F. Lhermitte, J. C. Gautier, R. Marteau and F. Chain (1963): Troubles de la conscience et mutisme akinétique. *Revue neurologique* 109: 115—131.
- A. Liberman (1957): Some Results of Research on Speech Perception. *Journal of the Acoustical Society of America* 29: 117—123.
- L. Stein (1925): Das universelle Stammeln im Lichte der vergleichenden Sprachwissenschaft. *Zeitschrift für die gesamte Neurologie und Psychiatrie* 95: 100—107.

DISCUSSION

Fischer-Jørgensen:

The cases of grown-up aphasics are rather difficult to judge; the fact that children may be able to perceive differences (e.g. t : k) before being able to produce them seems more conclusive; for this must mean that the coordination between perception and motor commands has not yet been established.

Fry:

As far as I know, no one would say that there is only one way of perceiving speech, namely by reference to the motor patterns by which we generate speech. The question is really: "In normal conditions, what contribution can be made by the motor system to the perception and reception of speech?" Most of the cases cited by Dr. Lebrun are not very relevant because the patients have at some time developed the motor mechanism of speech and there is no reason why the cortical patterns should not be available. The only pertinent cases are those in which the movements of speech have never been made and these simply serve to emphasize the fact that by the very nature of speech, we can never say that a certain operation is carried out in one way only and in no other. Experience shows that if such statements are made they are pretty sure to be wrong.

NOTICING WORD-BOUNDARIES: A BRIEF INVESTIGATION

W. R. LEE

The investigation was conducted in two stages. In the first stage ten speakers each recorded phrases such as *mice kill* and *my skill*. In the second stage ten listeners tried to identify what had been said. Both speakers and listeners spoke English as their first language.

The list of phrases was as follows: *Joy sprints*/*Joyce prints*, *pea-stalks*/*peace-talks*, *Park Road*/*Pa crowed*, *once tried*/*one stride*, *so dry*/*sowed rye*, *mice kill*/*my skill*, *Pa cried*/*Park Ride*, *Nye speaks*/*nice peaks*, *Lu slips*/*loose lips*, *house trained*/*how strained*, *race miles*/*Ray smiles*, *false printing*/*fall sprinting*, *fear strumpets*/*fierce trumpets*, *Joyce trips*/*Joy strips*.

The phrases were of more than one type. First, there were those in which the word-boundary occurs either between a vowel and two consonants (as in *my skill*) or between a consonant and a consonant (as in *mice kill*): there were sixteen of these V-CC and VC-C phrases. Secondly, there were those in which the boundary occurs either between a vowel and three consonants (as in *how strained*) or between a consonant and two or three consonants (as in *house trained*): there were eight such V-CCC or VC-CC phrases. Thirdly, there were four phrases belonging to the somewhat similar types VC-CC(C) or VCC-CC.¹ All the phrases chosen can stand on their own, with the possible exception of *race miles*. All but the pair *pea-stalks* and *peace-talks* appear to invite the use of two strong stresses. In choosing the phrases, no attempt was made systematically to include various types of sound; but eleven different vowels and a number of consonant combinations are represented.

The phrases were not presented (to either the speakers or the listeners) in sentences, but as isolates, the reason being that it was easier, so to speak, to equate them in this way, whereas to devise sentences in which each member of the pair occurred in identical contexts of stress and intonation would be extremely difficult and also perhaps unnecessary, since for the second stage of the investigation the context had to be eliminated, the aim of the investigation being to find whether the word boundaries were rightly perceived in the absence of contextual guidance.

The ten speakers, who were not of course told what was involved, were all women of 18 to 20 years old (except for one aged 45). They were speech-therapy students who

¹ Occurrences of phrases belonging to the second and third types seem to be fairly rare.

had not received instruction in juncture or word-division, though they had some knowledge of phonetics. They came from various parts of England and some spoke with traces of regional dialect.

Two lists were prepared, each containing 40 phrases. In each list, 14 of these phrases belonged to the selected pairs, no phrase occurring twice. The types were mixed; thus there were five V-CC phrases on Sheet 1 and three on Sheet 2, and so on. Besides these 28 phrases belonging to the material under observation, there were 52 other phrases (26 in each list) intended to distract the speakers from realising what the investigation was about. On the whole this 'blind' material consisted of phrases, such as *black smoke*, *mutton chop*, and *no time*, having roughly the same rhythmic character as the rest. Many of them contained also such features as incomplete and nasal plosion. The distractive material was successful in its purpose, and none of the speakers guessed the nature of the investigation.

Each speaker read from the lists individually, out of sight and hearing of the others, and was tape-recorded at 3.5 ins. (9.5 cms.) per second. Each speaker was separately asked beforehand not to make any effort to read distinctly, and for each a sample speed was given by reading two or three phrases not on the list. No opportunity was allowed of looking through the phrases in advance of recording. On the second occasion the speakers followed one another in a somewhat different order from the first.

Check sheets were then prepared for the ten listeners: men and women teachers in training (average age, 24) at another institution.² Each sheet bore the 28 phrases, arranged in numbered pairs, and no other material. The listeners were asked to listen to the tapes and to place a tick on their sheets against the phrase they thought they heard as each relevant item came along (the numbers had also been recorded). The relevant phrases had not been cut out and made into a new tape; instead they were brought into prominence by means of the volume control, which was turned down for the other items. This proved to be a speedy method of presenting the phrases to the listeners, though it had the disadvantage that some irrelevant material could be faintly heard, owing to the practical difficulty of turning down the volume quickly enough.

Since 28 phrases spoken by ten different speakers were listened to by ten listeners, 2800 judgements were involved. Of these, 79.7% were correct, the four men's and the six women's performances being on average about equally good. The least successful performer scored 67.1%,³ the most successful 90%.⁴ The least accurately

² I am grateful to the Phonetics Department of University College, London (and particularly to Professor A. C. Gimson and Mr J. D. O'Connor) and also to the Division of Language Teaching at the Institute of Education, London (and especially to Professor B. Pattison and Mr G. Broughton) for their provision of the facilities which enabled this investigation to take place.

³ Although this listener's first language was English, she was of mixed parentage.

⁴ In a similar investigation, concerned with other types of word-boundary and involving 2400 judgements, J. D. O'Connor and O. Tooley observed a lower degree of success in identification — 66.8%. See 'In Honour of Daniel Jones', 1964, p. 171.

identified phrase was *race miles* (51 errors), followed by *sowed rye* (41), *Pa crowed* (36), *nice peaks* (35), and *Lu slips* (35). The phrases most consistently identified with success were *one stride* and *Nye speaks* (2 errors each), followed by *peace-talks* and *house trained* (5 each) and *Park Road* (7). One is tempted to think that frequency of everyday occurrence may have had something to do with the ease with which a phrase was identified, but there is nothing in the evidence which clearly bears this out: one can hardly believe that *Nye speaks*, which was only misheard twice, is any commoner than *nice peaks*, which was misheard 35 times, though it is probably true that *Park Road* (7 errors) is more often to be met with than *Pa crowed* (36 errors).

Every listener except one listened more successfully to the second tape than to the first, the average improvement being slight (roughly 1.3 extra identifications out of 10; range, 0.5 to 2.6). Most improvement was shown by the listener with the lowest score. The improvement suggests that increasing practice in attempting to distinguish such word-boundaries may lead to increased skill; but the evidence is not of course enough to be conclusive.

Discrimination between phrases of the V-CCC and VC-CC(C) types, and also between those of the VC-CC(C) and VCC-CC types, was more accurate than discrimination between the V-CC and VC-C types. There were, however, only eight phrases belonging to the former pair of types, and four to the latter; whereas there were sixteen phrases belonging to the V-CC or VC-C types. The average number of mishearings of the phrases in this category was 45.

Initial aspiration of stops might perhaps be expected to give listeners a reliable clue. *Nice peaks*, however, was among the phrases most often misheard, though there was little mishearing of *peace-talks*.

Further analysis of the data obtained would be possible. It seems at the moment illegitimate to draw any conclusion except the broad one that *without guidance from context* it is hard even for native British listeners to recognize certain word boundaries in British English. (This conclusion does not necessarily apply to American English or to other types of word boundary.) One wonders whether it is worth while, in teaching the pronunciation of English, to spend a lot of time on drilling this type of feature.

DISCUSSION

Hill:

Professor Lee would find it profitable, I believe, to consult the elegant experiment by Prof. Lehiste (*Phonetica* vol. 5) if he has not done so. In my own teaching, I find it now clearer to stick to "juncture" as a syllable divides, as in *at all: a tall: atoll*. When there is no separator the division is in the (+), otherwise before or after it.

Krech:

Wurde bei Ihrer Untersuchungen eine physikalisch-akustische Analyse der Wortpaare durchgeführt, und wenn ja, welche Allophone der Vokale oder Konsonanten konnten in der Randstellung beobachtet werden? Es wäre wesentlich zu erfahren, welche Merkmale für die Perzeption relevant sind.

Lee:

ad *Krech*: No analysis of this kind was carried out, but it would of course be a useful extension of the research.

DIPHTHONGS VERSUS VOWEL SEQUENCES IN ESTONIAN

ILSE LEHISTE

This paper deals with the problem of distinguishing between diphthongs and vowel sequences containing a word boundary. The language in which the phenomenon has been studied is Estonian. The paper is based on acoustic analysis of 500 short sentences produced by one informant (the author). It is to be understood that generalizations drawn from this limited material serve only to set up working hypotheses which will be tested by analysis of a larger corpus of utterances, produced by several additional speakers.

There are nine vowels and a large number of diphthongs in the language. Of the nine vowels, [a e i u] may occur in stressed as well as unstressed position in any syllable; they may also occur as first or second element of a diphthong. The following combinations of these four vowels were studied: the diphthongs [iu], [ei], [ea], [eu], [ai], [ae], [au], and [ui], and the vowel sequences [i] + [u], [e] + [i], [e] + [a], [e] + [u], [a] + [i], [a] + [e], [a] + [u], and [u] + [i]. It was hypothesized that in a diphthong, V_1 would be similar in phonetic quality to a stressed short vowel, and V_2 to an unstressed vowel occurring in the second syllable of a disyllabic word. It was further assumed that the diphthongal sequence would differ from a sequence of the same vowels containing a word boundary in the relative stress of the components: in the diphthong, V_1 would be stressed, V_2 unstressed, whereas in the sequence containing a boundary, V_1 would be unstressed and V_2 stressed.

Broad-band spectrograms were made of all utterances. The sentences were also processed through an intensity meter and pitch meter (produced by B. Frøkjaer—Jensen, Engineering Firm of Electronic Instruments, Copenhagen) and displayed on a four-channel Mingograph (Mingograph No. 42—EM/122, Elema-Schönander, Stockholm). The results of the study are summarized in a series of acoustical vowel diagrams which were constructed on the basis of averages calculated from measurements made of broad-band spectrograms. Corresponding tables are likewise presented.

The first figure shows F_1 — F_2 positions of the stressed short vowels [i e a u] and positions of the same vowels forming the first element of a diphthong. The positions of stressed long and overlong vowels have likewise been plotted on the diagram. The figure shows that the phonetic values of stressed short vowels are remarkably similar to the first components of diphthongs. Both differ markedly from long and overlong vowels, which are phonetically close to each other.

Figure 2 compares the second components of diphthongs with target positions of stressed monophthongs. In the case of [i] and [u], the second components of overlong diphthongs are phonetically similar to long vowels; but with [a] and especially

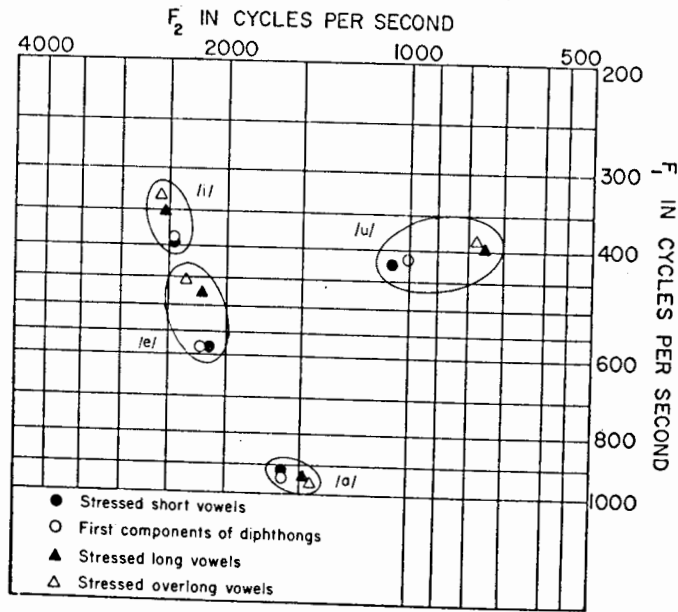


Figure 1. Acoustical vowel diagram representing $F_1 - F_2$ positions of the stressed vowels [i e a u] occurring in three quantities as monophthongs and as first components of diphthongs.

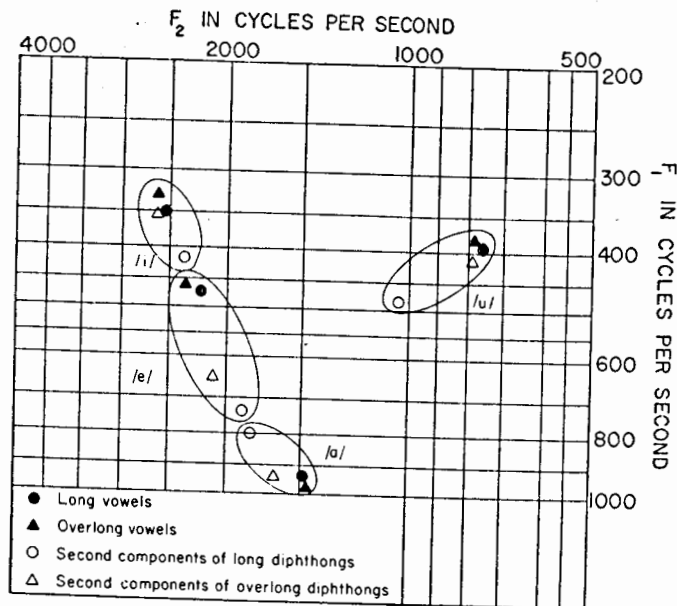


Figure 2. Acoustical vowel diagram representing $F_1 - F_2$ positions of long and overlong [i e a u] compared with the same vowels occurring as second components of long and overlong diphthongs.

with [e], no such statement can be made. The second components of long diphthongs are in each case distinctly different from stressed monophthongs.

Figure 3 compares V_2 of diphthongs with unstressed vowels in the second syllable of disyllabic words. The results are contradictory: for [u], unstressed vowels are similar in phonetic quality to V_2 of long diphthongs, but for [e], [a], and [i], the unstressed vowels are more similar to V_2 of overlong diphthongs.

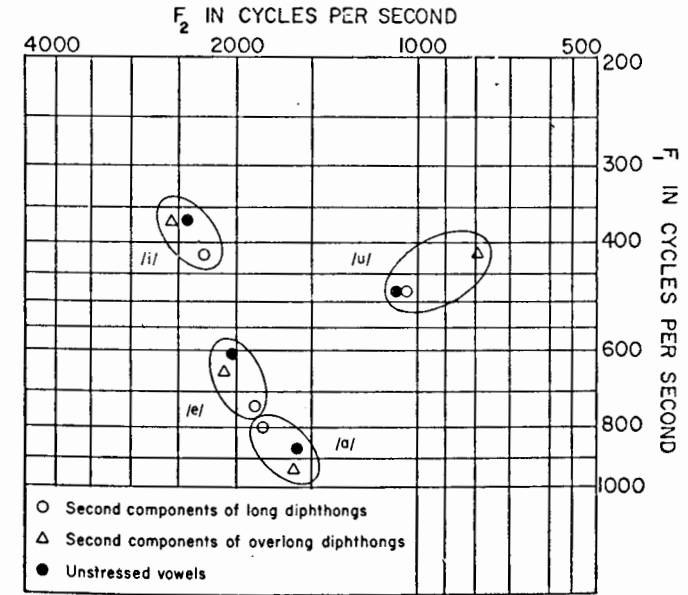


Figure 3. Acoustical vowel diagram representing $F_1 - F_2$ positions of [i e a u] occurring as second components of long and overlong diphthongs and as unstressed vowels in the second syllable of disyllabic words.

It seems that the second hypothesis cannot be verified on the basis of this material: the phonetic quality of the second component of diphthongs is not identical with the phonetic quality of unstressed vowels. Neither is it identical with that of stressed vowels, except for [i] and [u], where V_2 of overlong diphthongs was similar in quality to long and overlong monophthongal [i] and [u].

The clues provided by duration were considered next. The duration of the two segments was measured, assuming a boundary in the middle of the transition from the first to the second component. The two components of a diphthong were found to be almost equal in duration both in long and in overlong diphthongs: the lengthening of an overlong diphthong was apparently accomplished by a proportional lengthening of both components. The average durations of V_1 and V_2 of a long diphthong were 9.4 and 7.4 csec; in an overlong diphthong, the values were 13.2 and 13.9 csec. The average durations of vowels in a sequence containing a word boundary were 8.6 and 8.3 sec. Duration thus distinguishes a $V + V$ sequence from overlong diph-

thongs, but not from long diphthongs. The contribution of intensity to the difference between diphthongs and vowel sequences containing a word boundary came next under consideration. Only qualitative observations are available, since the test material was not suitable for a quantitative treatment of the data. Limitations of space make it impossible to present more than a single example.

Figure 4 contains an oscillogram, a fundamental frequency curve, and an intensity curve of the two utterances *Imad on kuivad* ("The days are dry") and *Seleks kulus terve eluiga* ("It took a whole lifetime"). The comparison is between the long diphthong [u \ddot{i}] in *kuivad* and the [u] + [i] sequence in *eluiga*, which contains a word boundary between [u] and [i]. The intensity curve shows an additional peak during the second component of the [u] + [i] sequence *eluiga*. Such separate peaks were frequently found in V + V sequences; they did not occur in diphthongs.

In summary, a complex set of distinctions was found between diphthongs and V + V sequences. The difference between long diphthongs and V + V sequences was partly due to the difference in phonetic quality between the second component

Table I. Comparison of stressed monophthongs with the first component of diphthongs.

Vowel	Number of occurrences	Average Duration in csec	Formant frequencies in cycles per sec.		
			F ₁	F ₂	F ₃
Short [i]	15	8.6	395	2495	3115
Long [i]	4	14.7	350	2540	3200
Overlong [i]	12	19.4	325	2560	3320
[i] as V ₁ *	2/3	11.5/14.0	380	2475	3035
Short [e]	13	9.6	585	2130	2940
Long [e]	7	12.1	470	2205	2885
Overlong [e]	12	18.3	460	2350	2985
[e] as V ₁	11/16	8.4/12.9	580	2190	2975
Short [a]	21	11.4	925	1600	2695
Long [a]	8	15.9	950	1495	2720
Overlong [a]	23	22.1	975	1445	2765
[a] as V ₁	25/18	10.2/13.9	955	1605	2705
Short [u]	10	9.7	420	1060	2780
Long [u]	14	17.2	395	750	2855
Overlong [u]	26	21.0	385	770	2845
[u] as V ₁	5/4	6.8/11.3	410	1000	2730

* The formant positions for the first component of a diphthong are averages for both long and overlong diphthongs. The number of each type and the durations are given separately; the first number refers to the first segment of long diphthongs, the second to the first segment of overlong diphthongs.

of diphthongs and between stressed vowels. The difference between V + V sequences and overlong diphthongs was primarily durational. Intensity differences provided an additional clue that seemed to be relatively independent of vowel quality. An

Table II. Second components of diphthongs.

Vowel	Second component of long diphthongs					Second component of overlong diphthongs				
	N	Dur.	F ₁	F ₂	F ₃	N	Dur.	F ₁	F ₂	F ₃
[i]	28	6.4	415	2310	3055	19	12.2	365	2560	3215
[e]	2	6.0	740	1865	2800	8	13.4	655	2060	2875
[a]	2	8.0	800	1815	2775	7	13.4	950	1620	2700
[u]	11	7.5	480	1030	2725	7	16.0	415	785	2845

Table III. Unstressed vowels in the second syllable of disyllabic words.

Vowel	Quantity of preceding syllable	Number of occurrences	Duration in csec	Formant positions in cycles per second		
				F ₁	F ₂	F ₃
[i]	Short	21	10.9	355	2520	3120
	Long	14	9.2	365	2470	3120
	Overlong	29	9.9	370	2465	3115
	Average:			365	2485	3120
[e]	Short	21	9.6	625	2030	2890
	Long	21	8.1	595	1980	2845
	Overlong	35	9.4	615	2005	2885
	Average:			610	2005	2875
[a]	Short	33	12.3	900	1530	2685
	Long	60	9.6	825	1615	2720
	Overlong	51	9.5	880	1625	2775
	Average:			870	1590	2725
[u]	Short	14	11.9	445	975	2810
	Long	9	9.7	505	1110	2815
	Overlong	21	8.9	470	1125	2735
	Average:			475	1070	2785

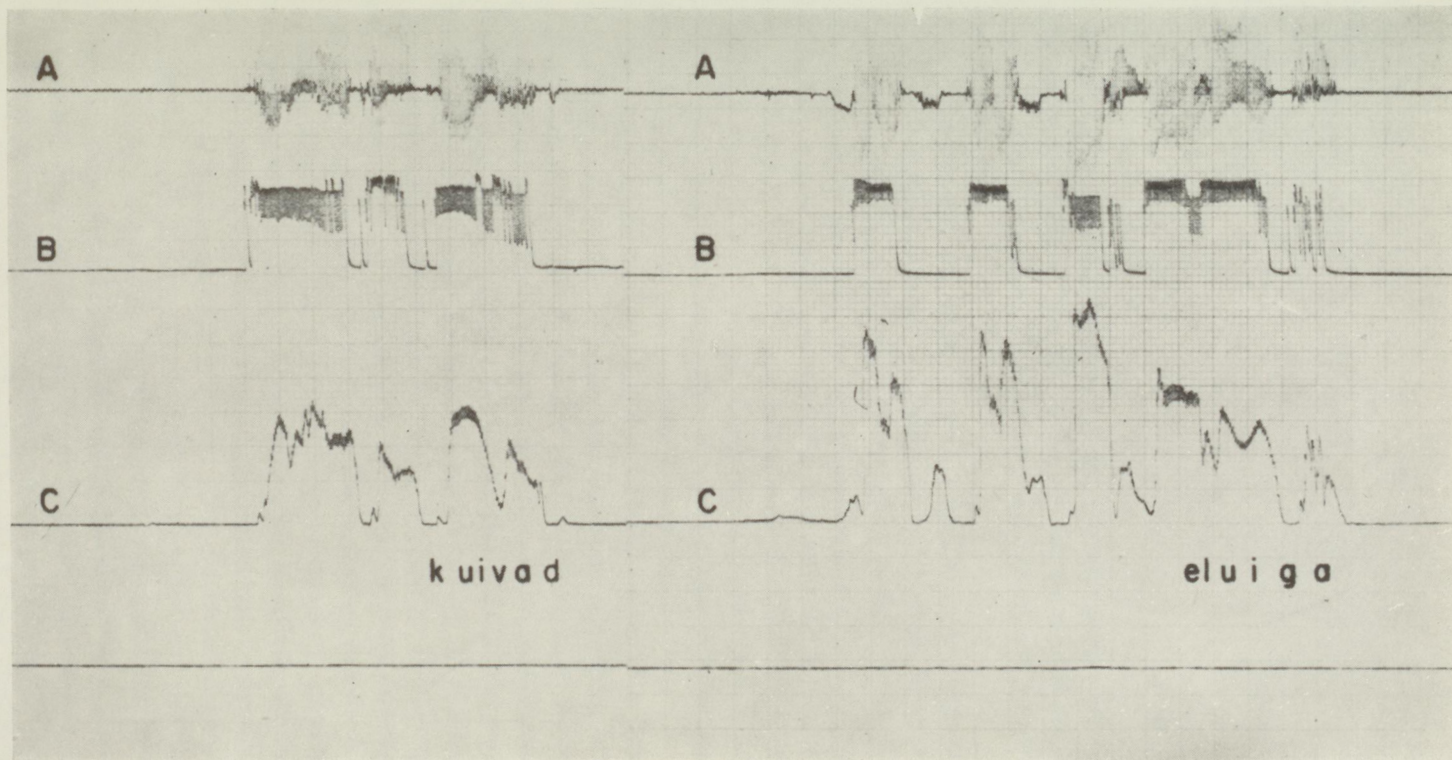
unexpected result was the discovery that the overlength of long diphthongs is distributed evenly over both components of the diphthong. This finding supports the assumption that the domain of overlength is the whole syllable rather than one of the segments of the diphthong.

DISCUSSION

Hint:

In the phonological pattern of Estonian the long vowels function as equivalent to long diphthongs and overlong vowels to overlong diphthongs. It is very interesting to hear now that they are similar in respect to their phonetic structure too, so far as quantity and duration are concerned. This conforms to phonological expectations.

- A. Oscillogram
- B. Fundamental frequency
- C. Intensity



1. Ilmad on kuivad

2. Selleks kulus terve eluga

Figure 4. Oscillogram, fundamental frequency curve and intensity curve of two utterances, one containing the long diphthong [ui], the other the sequence [u] + [i].

EINFLUSS DER MUNDART AUF DIE ARTIKULATION DER KINDER

VLADISLAV LEJSKA

Die Frequenz des Stammelns bei Schulkindern beträgt überall auf der Welt ungefähr 20 %. Den einheimischen Autoren gemäß werden am häufigsten die Schwinglaute R und Ř betroffen, deren fehlerhafte Aussprache sich auch mit dem zunehmenden Alter am wenigsten bessert.

Wir führten eine Untersuchung der Aussprache bei Schülern der ersten Jahrgänge an Schulen in den Mundart-Gebieten von Hornácko und Podluží durch. Beide mundartlichen Gruppen gehören zum mährisch-slowakischen Dialekt. In Hornácko kommt im Dialekt der Laut Ř gar nicht vor und wird durch den Laut R ersetzt. In Podluží ersetzt man den Laut L durch drei verschiedene Laute und, zwar: a) durch einen harten, dem R ähnlichen Laut; b) einen weichen Gaumenlaut und c) einen umhüllten, den U-Laut akustisch sowie auch durch die Einstellung der Sprechorgane ähnlichen Laut. In diesen Gebieten sprechen Erwachsene und Kinder Dialekt. Schulkinder gebrauchen die Schriftsprache nur beim Unterricht.

Wir haben die Ergebnisse der Durchforschung der Artikulation mit dem Zustand der Entwicklung der Artikulation bei Kindern aus Gebieten verglichen, wo die Umgangssprache der Bewohner praktisch von der Schriftsprache nicht abweicht.

Im ganzen haben wir bei 1056 Kindern (49,5 % Knaben; 50,5 % Mädchen) eine diesbezügliche Untersuchung durchgeführt. Zu Beginn des Schuljahres war die Frequenz der Unfähigkeit den Ř-Laut zu artikulieren im Gebiet Hornácko fast 100 %, aber in der kontrollierten Gruppe nur bei 40 % der Kinder. (Die Frequenz der Rhotazismen, d. h. nur R-Laut, bewegte sich in beiden Gruppen zwischen 16 bis 19 %). Am Schuljahrschluß stellten wir im Gebiet Hornácko fest, daß die Kinder spontan, ohne fachmännische Hilfe den Ř-Laut richtig aussprachen, obwohl die Entwicklung der richtigen Aussprache nur durch die Schule beeinflusst wurde. Die Kinder gebrauchen den Schwinglaut Ř wieder nur beim Unterricht. In der kontrollierten Gruppe hat sich die Aussprache des Ř nur bei 25 % der Dyslalier gebessert. In Gebieten mit dialektisch artikuliertem L stellten wir dagegen fest, daß sich die drei dort vorkommenden Typen des L-Lautes im Vergleich mit dem Ř-Laut nicht so leicht und schnell bessern. Sogar Schüler der 5. Jahrgänge leiden an diesen „Lambda-zismus“ und bei manchen bleibt er bestehen. Diese Tatsache können wir uns so erklären, daß sich der akustisch-phonetische Reflex weit besser bei dem Übergang vom R-Laut zum Ř-Laut bewährt, wogegen der dialektische L-Laut akustisch von

dem schriftsprachlichen wenig abweicht, so daß die Erlangung der richtigen Aussprache nicht so sehr vom Gehör unterstützt zu werden braucht. Bei den Rhotazismen wurde immer der R-Laut durch den im Dialekt üblichen L-Laut ersetzt.

Bei der Schlußbewertung unserer Aktion haben wir uns einige Fragen gestellt. Wenn auch die Antworten überzeugend klingen mögen, so sind sie vorläufig nur als Hypothesen zu werten.

Erste Frage:

Wie ist es zu erklären, daß fast alle Kinder aus den betreffenden Gebieten, welche vor dem Schulbesuch den Ř-Laut gar nicht artikulieren konnten, binnen einem Jahr den in der tschechischen Sprache schwierigsten Laut leicht artikulieren lernen? Es ist doch ein großer Unterschied zwischen der vom frühesten Kindesalter an bestehenden Stimulation und jener beschränkten Stimulation in unseren Fällen. Vorläufig haben wir drei Faktoren festgestellt, die auf Grund mehrjähriger persönlicher Erkenntnisse in betreffenden Gebieten entstanden sind und von denen wir überzeugt sind, daß sie einen starken Einfluß auf die Dynamik der Sprache haben könnten.

1. Ein angeborenes Talent für Musik und Gesang, das weitgehend gepflegt wird und ein Charakterzug der Bewohner des Hornácko-Gebietes ist. Ihre Lieder sind melodisch sehr anspruchsvoll, so daß die Qualität des Gehörs auf der Höhe ist.

2. Eine vollkommene Vibration des R-Lautes ist dort schon fixiert, wo ihn anderswo die Kinder noch durch den L-Laut ersetzen. So kommt es, daß der Übergang zum Ř erleichtert ist.

3. Die erhöhte Aufmerksamkeit, welche die Lehrer der 1. Jahrgänge unwillkürlich der Aussprache des Ř-Lautes widmen.

Zweite Frage:

Kann diese dialektische Aussprache als Dyslalie betrachtet werden? Sicher nicht! Denn das, was in einem anderen Gebiet als Dyslalie angenommen wird, sollte nicht als solche in denjenigen Gebieten gelten, wo diese, von der Schriftsprache abweichende Aussprache von allen Bewohnern normal gebraucht wird. Als Dyslalie kann diese abweichende Aussprache erst beim Eintritt in die Schule, und zwar nur während des Unterrichtes gelten. In diesem Falle sollten wir die Artikulation von einem anderen logopaedischen Standpunkt betrachten.

Dritte Frage:

Diese beiden Probleme hängen sehr eng mit der praktischen Schlußfolgerung in bezug auf die logopaedische Arbeit zusammen. Die logopaedische Tätigkeit ist hier nämlich nicht dieselbe wie in anderen Gebieten. Es ist zwingend notwendig, daß der Logopäde vorher die Abweichungen von der Schriftsprache kennen lernt. Wir

nehmen an — und die Untersuchungen haben es uns restlos bestätigt —, daß es unnötig ist die abweichende Artikulation der Ř- und L-Laute schon im Vorschulalter unbedingt bessern zu wollen. Die hier erforderliche Spracherziehung wird doch mit großem Effekt von der Schule in einem passenden Milieu spontan durchgeführt.

DISCUSSION

Martens:

Übergang von Mundart auf Hochsprache bedeutet — genau wie das Erlernen einer Fremdsprache — den Wechsel von einem phonologischen System zu einem anderen. Fehlleistungen, die dabei auftreten, kann man selbstverständlich nicht als Sprechfehler klassifizieren. Dennoch können in solchen Fällen logopädische Methoden oft schneller zum Ziel führen. Dies zeigt, daß alle Sprachlehrer großen Nutzen aus Kenntnissen auf dem Gebiet der Sprachpathologie und der Sprachheilpädagogik ziehen können.

Sovijärvi:

Drei verschiedene (l)-Laute kommen auch in den finnischen Dialekten vor, nämlich, das neutrale, palatalisierte und leicht velarisierte (l). Die zwei letzten Abweichungen von dem normalen bleiben manchmal durch das ganze Leben in dem Sprechen der geschulten Leute, aber sind nicht als Dyslalien betrachtet. Wenn das Kind seine Zungenmuskulatur gemäß der betreffenden Übungsteste beherrscht, spricht man bei uns von sprechtechnischen oder digressiven Lautfehlern. Nur die transgressiven Lautfehler sind funktionelle oder andere Dyslalien.

SOME NOTES ON THE EXPERIMENTAL STUDY OF SPECTRAL PARAMETERS OF VOWEL-CONSONANT TRANSITIONS IN ESTONIAN

GEORG LIIV*

The aim of this spectrographic study was to examine certain dynamic aspects of vowel articulation. The majority of studies on acoustic characteristics of vowels present measurements only of formant frequencies at fixed points, such as points located centrally within vowels or points at the boundaries between vowels and consonants, and hence provide only a partial description of the vowel. This study endeavours to analyze the continuous change of the whole acoustic structure throughout the entire vocalic nucleus of the syllable.

The spoken samples to be analyzed were drawn from real words of Standard Estonian and had the form CVC, mainly in phonemically asymmetrical syllables. Nine stressed vowels /*i e ü a o u ü ö õ* = [ē]/ embedded in 14 initial and 15 final consonants that can appear in Estonian syllables were used, mainly in such word structures as /CVC₁::(C₂) (V) (C)/, /CV::C/, /CV::C::/, or /CV(:) (:) C(:) (:) V/ (the first syllable vowel or intervocalic consonant occurring in short, long or overlong distinctive quantities). The vowels /*i a u*/ were studied in all possible combinations, the others being analyzed less systematically. The whole speech inventory consisted of 237 words uttered by four adult male and two female informants.

The conclusions are drawn from the synchronous analysis of the dynamic sound spectrograms and sections (film speed 64 frames/sec) obtained by means of a high-speed 52-channel dynamic sound spectrograph of the Experimental Phonetics Laboratory of the Language and Literature Institute of the Estonian S. S. R. Academy of Sciences (the synchronous analysis of sections and dynamic spectrograms was guaranteed by indicating the time locations of sections at the bottom of the dynamic spectrograms as well as by the automatic enumeration of the sections with frame counters whose readings were registered on both films), and from the segmentation of the vowels with a gating circuit on the basis of auditory analysis.

The frequencies, intensities and bandwidths of the four lowest formants as a function of time were measured, and target positions relative to the undershoot phenomena and terminal values of vowel formants were described. Special attention

* From the Experimental Phonetics Laboratory, Institute of Language and Literature, Academy of Sciences of the Estonian S.S.R.

was paid to the temporal characteristics of the coarticulatory acoustic structures and the manner in which the terminal spectral values are approached and left by the formants during transitional segments. A description of the dynamic aspects in each vowel was based on approximately 330 measurements and it presents several additional parameters: extent and tempo of frequency change as transition curvature characteristics; total vowel duration; durations of the so-called quasi-stationary segment, initial and final transitions; duration of the individual formants, the points of their on- and offset, etc. As a by-product of the study, data on spectral irregularities were examined. It should be noted that the principal trends and durations of changes in acoustic structure were determined on the basis of dynamic spectrograms, whereas all spectral parameters were measured from sections yielding information for spectral samples spaced 16 msec apart throughout these vocalic portions.

In interpreting these acoustic data, it was found expedient to examine to some extent the nature and dynamic constraints of the articulatory activities producing these time-varying spectral parameters on the basis of cineradiographic films synchronized with sound spectrograms. To ensure the highest possible signal/noise ratio during simultaneous high-quality sound recording, the fluorographic equipment

Dynamic spectrogram of the Estonian word /'põl:ist/ (Elativ Sg. of the place-name *Põlli*) with some frames of the corresponding cineradiographic film (speed 50 frames/sec). The time locations of synchronously filmed sections are registered by means of small black marks along the lower margin of the dynamic spectrogram (speed 64 frames/sec). Vertical markings in the upper part of the spectrogram designate time intervals: the distance between shorter markings = 20 msec, that between longer markings = 100 msec. Small horizontal lines on the upper edge of the spectrogram register exposures (10 msec) of cineradiographic frames (to facilitate enumeration every first and tenth frame is marked with a darker and slightly heavier line; e.g. No. 39). Vertical arrows with numbers at the top of the spectrogram refer to the numerical order of the frames corresponding to certain time intervals. The same numbers also occur in the lower left-hand corner of the frames. To facilitate the counting of X-ray frames, every frame is marked with a light patch in the lower right-hand corner; every first and tenth frame also has an additional patch in the upper right-hand corner, e.g. No. 39.

The cineradiographic frames reproduced in the Figure show the articulatory movements from the "culminational" phase (No. 34) of /o/ through terminal transitions towards [i]-ness (which is the most essential acoustic correlate of palatalization in Estonian) up to the production of sharpened /i/ (No. 39). The continuous shift of the narrowest stricture in the vocal tract towards the front part of the oral cavity can be clearly observed (formed between the back of the tongue and the velum in Frame No. 34; between the blade and the alveoli with a conspicuous decrease in its cross-sectional area in No. 38). At the same time the width of the pharynx increases considerably, the valleculae epiglottidis widen, and there is some rise of the glottis. The continuously changing configuration of the vocal tract produces a very noticeable terminal transition (duration 95 msec out of a total vocalic duration of 140 msec). Note, for instance, the steady movement of F_2 from 975 cps (time interval corresponding to Frame No. 34) to 2475 cps (No. 38), and the opposite shift of F_1 from 750 cps to 525 cps. High female voice.

The Figure provides striking evidence of the lack of any steady state portion whatever in the production of the syllabic nucleus.

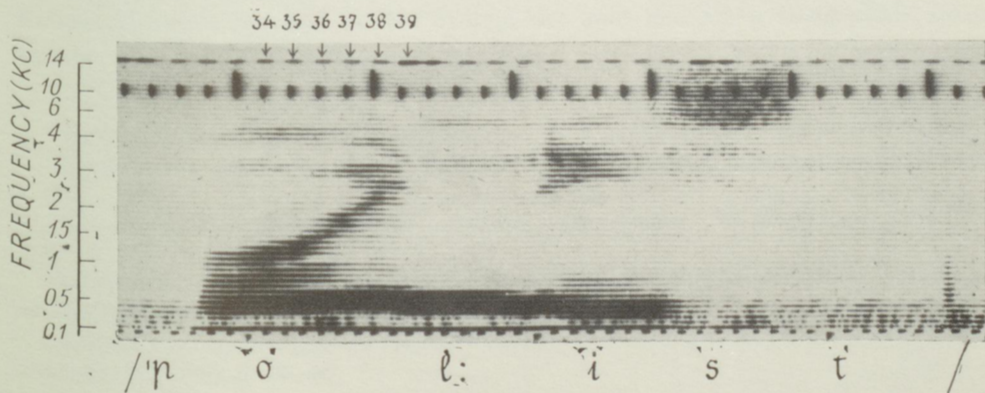
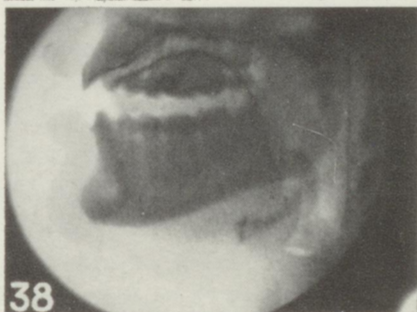
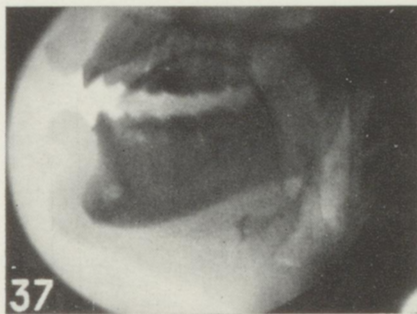
(TuR D 1000-2 six-valve roentgen apparatus, 5" Philips image intensifier, 35 mm Arriflex film camera) was placed in two adjoining insulated rooms (image intensifier input and informant in one room, the fluorescent screen of intensifier in another): the X-ray tube and collimator were padded with felt, a directional microphone was used, curtains of thick cloth were hung about the informant. Fine results have been achieved in minimizing and standardizing the distortion and enlargement resulting from the conical spread of roentgen rays (registered on film by means of a metal grid). For this purpose the head posture (midsagittal plane) of the informant was fixed by a special head holder in a position strictly parallel to the vertical plane of the image intensifier tube and orthogonal to the central ray, the input surface of the image intensifier being brought into immediate contact with the informant's anatomical structure to be analyzed; the X-ray tube was moved to a distance of 2.5 metres from the midsagittal plane of the informant. Film speed 50 frames/sec. A special synchronizing device was employed to register X-ray frame exposures on the upper edge of the dynamic spectrogram (see Fig.).

In analyzing X-ray movies frame-by-frame, special attention was paid to the timing and interrelations of the movements of the various components of speech-generating structures, and to the associated acoustic outputs.

The complex techniques used in this study have provided useful data on the articulatory gestures and resulting consistent changes in acoustic structures of syllabic nuclei depending upon the place and manner of production of neighbouring consonants and upon inherent acoustic properties of the particular vowel. The temporal features of coarticulatory structures may be the result of mechanical and physiological inertia of the speech-generating system, and the result of the peculiarity of the control mechanism and overlap of neural instructions.

Work on the collection of further data is in progress. Some implications of the findings for a future dynamic theory of speech generation will be proposed.

The accompanying Figure exemplifies the materials obtained by means of the complex techniques used in this investigation.



DIE BEEINFLUSSUNG DER BEURTEILUNG ISOLIERTER SYNTHETISCHER VOKALE DURCH SUKZESSIVKONTRAST

GERHART LINDNER*

Die Decodierung ist ein komplizierter Prozeß der Informations- und Datenverarbeitung, der den Kommunikationsprozeß abschließt. Sie kann nicht als eine rein mechanistische Projektion der Außenwelt in das Bewußtsein angesehen werden.¹

Das Signal und das Ergebnis seiner Decodierung durch den Perzipienten sind über ein Gefüge von Bedingungen miteinander verbunden, wobei objektive und subjektive Bedingungen die Beurteilung eines lautsprachlichen Signals beeinflussen. Einige dieser Bedingungen der auditiven Beurteilung konnten herausgearbeitet werden.² Werden zu diesen Versuchen synthetisch erzeugte vokalähnliche Klänge verwendet, so liegen besonders günstige Voraussetzungen vor, weil der Experimentator von individuellen Eigenarten, wie sie gesprochenen Vokalen immer anhaften, unabhängig ist. Unter geeigneten Versuchsbedingungen ist es möglich, individuelle, von der Zugehörigkeit zu einer Dialektgruppe geprägte und vom Kontext abhängige Bedingungen nachzuweisen.

Die Veränderung der Beurteilung von Vokalen durch den Kontrast zu anderen Vokalen soll im folgenden im Mittelpunkt stehen. Der Sukzessivkontrast zweier Vokale bewirkt eine Verschiebung der Grenzen zwischen den Vokalfeldern. Um seine Wirkung möglichst unverfälscht zu untersuchen, wurde eine Anordnung verwendet, bei der störende Nebenwirkungen weitgehend ausgeschaltet waren.³

In vorangegangenen Versuchen waren mit deutschsprachigen Perzipienten die Beurteilungsdaten für synthetisch erzeugte vierformantige Klänge und ihre Streuungen ermittelt worden.⁴ Die Versuchspersonen (Vpn) hatten dabei die Aufgabe, diese Klänge zu beurteilen und in das System der acht deutschen Langvokale einzuordnen.

* Humboldt-Universität zu Berlin (DDR).

¹ Klaus, G.: Die Macht des Wortes, VEB Deutscher Verlag der Wissenschaften, Berlin 1964, S. 17.

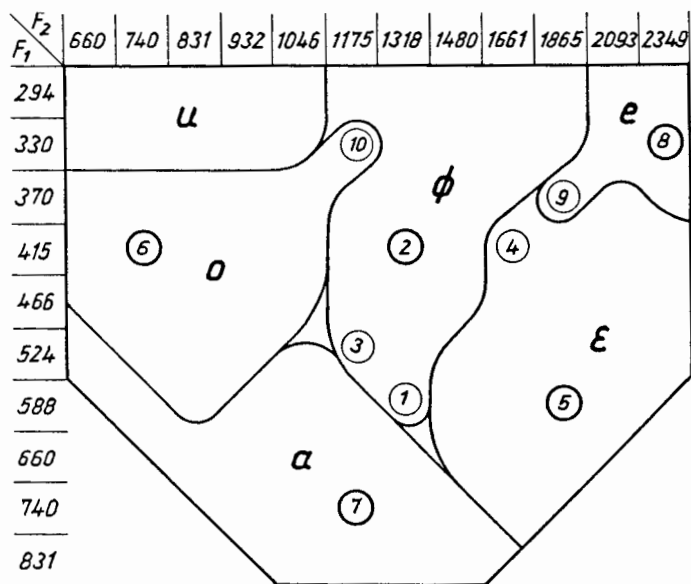
² Lindner, G.: Zur Bedingtheit der auditiven Wahrnehmung. *Die Sonderschule* 12 (1967), S. 3—14.

³ Lindner, G.: Veränderung der Beurteilung synthetischer Vokale unter dem Einfluß des Sukzessivkontrastes, *Ztschr. f. Phonetik* usw. 19 (1966), S. 287—307.

⁴ Lindner, G.: Beurteilung synthetischer Vokale durch deutschsprachige Hörer, *Ztschr. f. Phonetik* usw. 19 (1966), S. 45—90.

Für den Kontrasttest wurden für $F_0 = 110$ Hz und $F_0 = 220$ Hz je 10 Vokale ausgewählt. Fünf von ihnen hatten große interindividuelle Übereinstimmung. Sie werden als eindeutige Vokale bezeichnet. Sie liegen inmitten der Vokalfelder.

Weiterhin wurden fünf Grenzvokale verwendet, bei denen sich die Urteile der Perzipienten auf zwei oder drei der möglichen Urteile verteilen. Die Lage der ausgewählten Vokale in einem F_1/F_2 -Zuordnungs-Schema und in Relation zu den Vokalfeldern, die sich bei den vorangegangenen Versuchen ergeben hatten, zeigen Abb. 1 und 2.

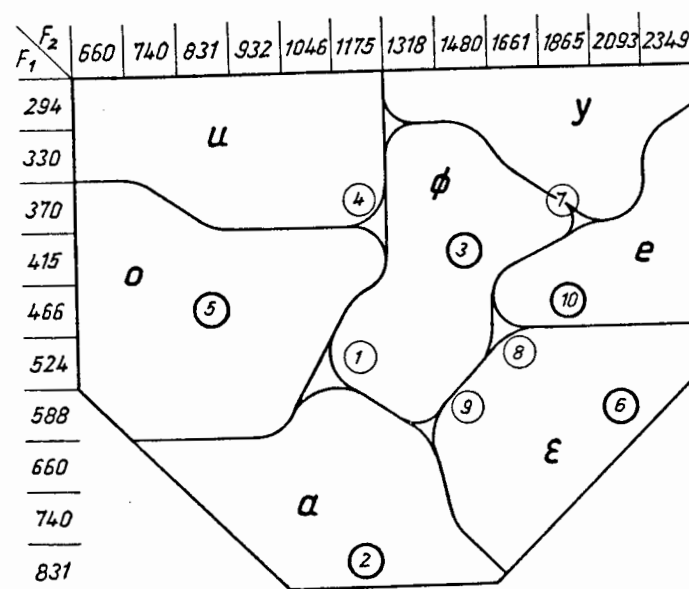


- eindeutige Vokale
- Mischvokale

Abb. 1. Lage der Vokale bei $F_0 = 110$ Hz in Relation zu den Beurteilungsfeldern im F_1/F_2 -Zuordnungs-Schema.

Die für jede Grundtonfrequenz ausgewählten Vokale wurden jeder mit jedem zu Zweiergruppen kombiniert. So entstehen für jeden der beiden Versuche 100 Paare, die wiederum von deutschsprachigen, normalhörenden Vpn beurteilt wurden. Die Vpn mußten erst die beiden, durch eine Pause von 1 s getrennten Vokale anhören und dann beide Urteile in ein Protokollblatt eintragen. Die Vokalpaare waren durch die Ansage und eine Pause von 5 s voneinander getrennt. Jeder Vokalklang mußte insgesamt 20mal beurteilt werden, und zwar unter verschiedenen Kontrastbedingungen. Die Ergebnisse der beiden Versuche mit verschiedenen Grundtönen stimmen gut überein.

Die Beurteilung der eindeutigen Vokale bleibt im wesentlichen konstant. Durch den Kontrast wird die Beurteilung nur in insgesamt 5% (8%) der Fälle verändert. Die Anzahl der Urteilsveränderungen liegt bei den Grenzvokalen erheblich höher bei 29% (25%). Die Angaben in Klammern beziehen sich auf den Versuch mit dem Grundton von 220 Hz.



- eindeutige Vokale
- Mischvokale

Abb. 2. Lage der Vokale bei $F_0 = 220$ Hz.

Die Besonderheiten der Kontrastwirkung lassen sich am klarsten erkennen, wenn ein Grenzvokal mit einem eindeutigen zu einem Paar zusammengefaßt wurde: Wenn der Grenzvokal über eine Valenz verfügt, die mit der des eindeutigen Vokals übereinstimmt oder ihm verwandt ist (z. B. a mit \tilde{a}), so wird im Kontrast diese Valenz erheblich vermindert. Diese Kontrastwirkung ist sowohl progressiv, d. h. wenn der Grenzvokal dem eindeutigen folgt, als auch regressiv, d. h. wenn der Grenzvokal dem eindeutigen vorangeht, vorhanden. Obwohl die regressive Kontrastwirkung etwa 10% schwächer ist als die progressive, ist ihr Nachweis außerordentlich wichtig; denn dadurch läßt sich der Kontrast eindeutig von der Adaptation unterscheiden (Abb. 3 und 4).

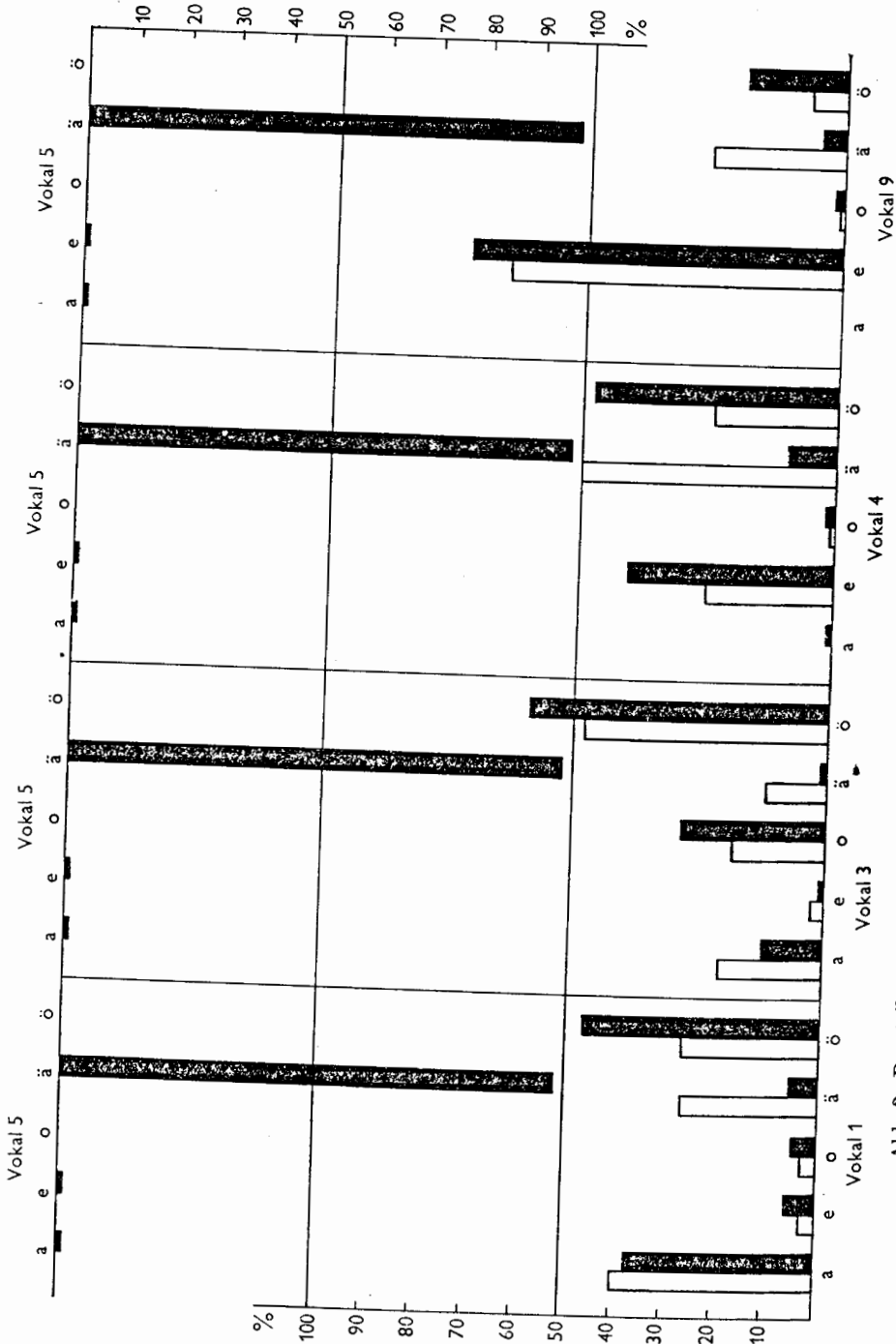


Abb. 3. Beurteilung verschiedener Grenzvokale im Kontrast mit dem gleichen eindeutigen Vokal.
Umrandete Säulen: Durchschnittswerte. Volle Säulen: Beurteilung im Kontrast.

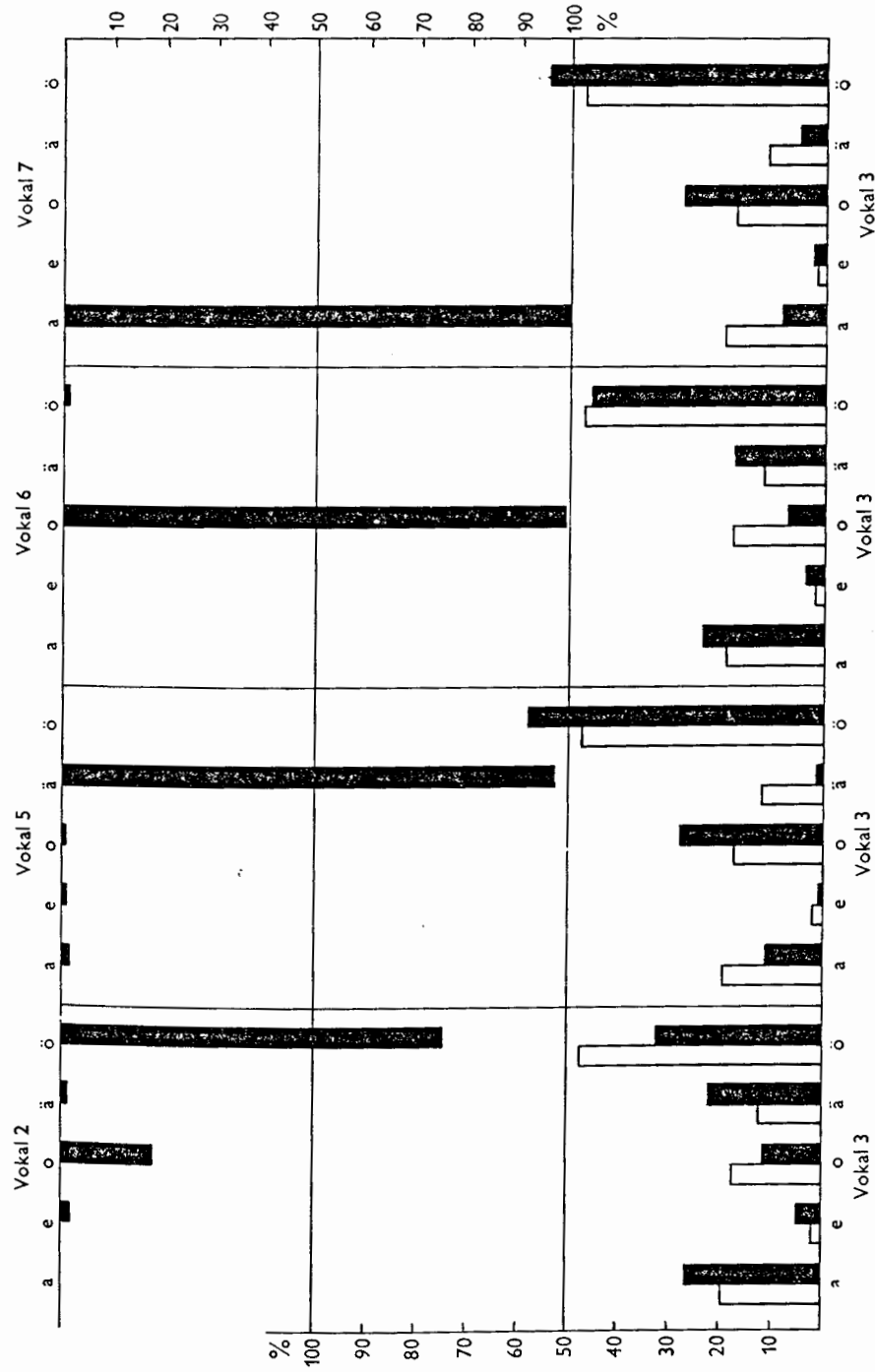


Abb. 4. Beurteilung desselben Grenzvokals im Kontrast mit verschiedenen eindeutigen Vokalen.

Wenn zwei Grenzvokale miteinander im Kontrast stehen, sind die gleichen Erscheinungen zu beobachten, nur daß sie gleichzeitig bei beiden Vokalen wirken. Dabei reduziert die dominante Valenz eines jeden Grenzvokals die Größe der gleichen Valenz des Kontrastvokals. Die verdrängten Urteile erhöhen die Größe der Dominante.

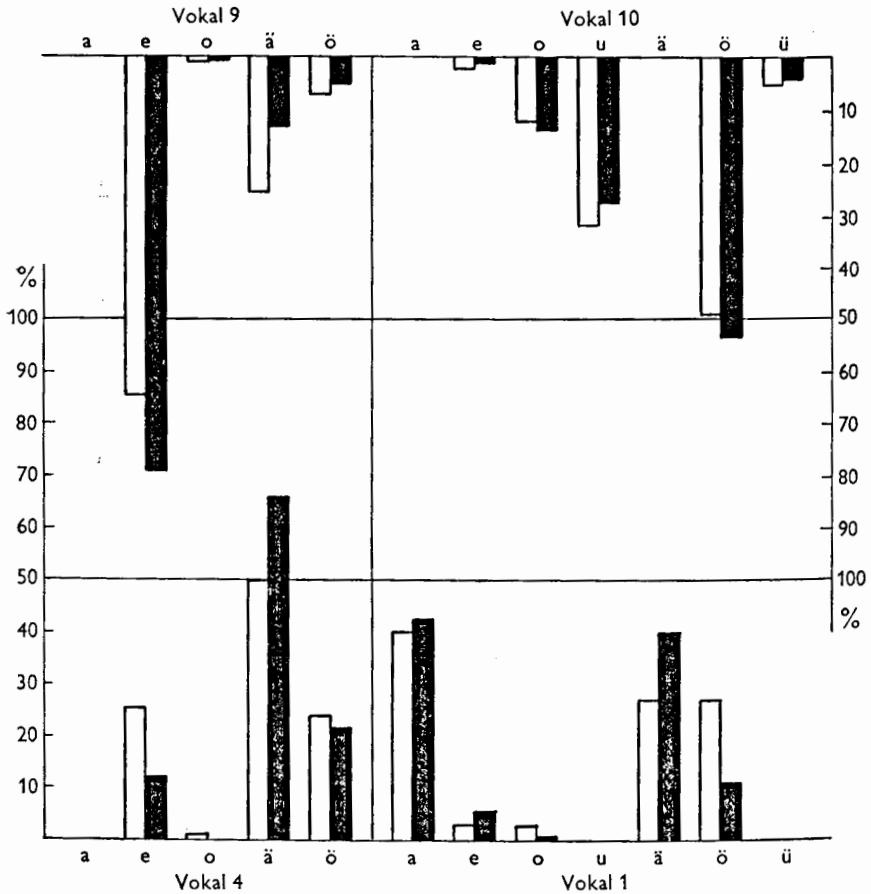


Abb. 5. Grenzvokale im gegenseitigen Kontrast.

Das bedeutet, daß beim Zusammentreffen zweier Grenzvokale die Dominanten beider stärker hervortreten. Mit anderen Worten: Beim gegenseitigen Kontrast zweier Grenzvokale werden ihre wesentlichen Merkmale eindeutiger (Abb. 5). Es ist wahrscheinlich, daß dieser Prozeß im mehrseitigen Kontrast noch weiter verstärkt wird.

Der Kontrast ist ein Bestandteil des Decodierungsprozesses. Dabei werden die erkennbaren Unterschiede zwischen einzelnen Signalsegmenten, die der Perzipient eindeutig bestimmt, stärker bewertet als die absolute Lage der Parameter dieser Signalsegmente.

LENGTH OF VOWELS IN SOTAK DIALECTS

JOZEF LIŠKA

The Sotak dialects are a small group of Eastern Slovak dialects in the North East corner of Czechoslovakia. While the rest of Eastern Slovak dialects have been more or less stabilized for some centuries now, so that they have no long vowels, and the stress is on the penultimate syllable, the Sotak group, even during the last seventy years has been undergoing important phonetic and phonological changes.

As this development is still going on we have only considered the speech of the oldest people. According to our findings, the length of the vowels is still relevant (phonemic) with the oldest speakers in the Sotak dialects.

Some authors, however, are not of the same opinion.

Samuel Czambel, for instance, the author of the monograph on the Eastern Slovak dialects published in 1906¹ and Zdzisław Stieber,² the Polish linguist (in the thirties before the Second World War) do not mention the long vowels in the Sotak dialects at all.

Phonetically, the absolute lengths of the Sotak vowels are variable, depending on their position in words, on the adjacent sounds, on the presence of stress and on pitch.

Our sonographs also show that both long and short vowels may be lengthened before a pause or at the end of a sentence due to the influence of intonation.

As even at present, some of the authors³ think the length of the Sotak vowels is the result of stress only, our primary task then, is to study the connection between the vowel length and stress, and how they influence each other.

Three facts have emerged from our investigation: First—The long vowels always have the main stress. Second—The short vowels can also occur under stress. Third—The lengthening of both long and short vowels when stressed is generally known, but the increase in duration is not noticeable and is not a significant feature.

In our opinion, the length of the Sotak vowels should be observed when they are in the same or similar positions, under the same stress, and with the same pitch. Our

¹ Czambel, S.: *Slovenská reč a jej miesto v rodine slovanských jazykov* (Turč. Sv. Martin 1906),

² Stieber, Z.: *Przyczynek do dialektologii Zemplina, Sborník Matice slovenskej XIII, 1935, 60—63.*

³ *Atlas slovenských nářečí* which is in print now has no long vowels in the Sotak area.

investigations show, that, in such conditions, the duration seems to be the principal and significant feature of contrast between long and short vowels, and is capable of distinguishing the meaning of words.

There are four pairs of long and short vowels in the Sotak vowel system at present. Let me introduce them in words and listen to them in sentences as recorded on magnetic tape (they are repeated twice each).

The first pair is [a]—[a:]: *para* “steam”—*pa:ra* “a pair”, *spa:l'i* “they slept”—*spa:l'i* “he will burn”.

The second pair is [æ]—[æ:] which occur mostly after palatal consonants: *p'æta* “a heel”—*p'æ:ta* “the fifth”, *p'æna* “a foam”—*p'æ:na* “tipsy”.













The third pair is [o]—[o:]: *pot* “sweat”, “under”—*po:t* “a loft”.

The fourth pair is [e]—[e:]: *ščep'u* “I graft”—*ščē:p'u* “I split”.

As to former long vowels [i:] and [u:] they do not differ in length from the corresponding short vowels any more.

According to our measurements on the Sona-Graph and oscillograph, the length of the individual vowel pairs is as shown on the Table I and Table II.

Table I (Subject A is M. B., Dlhé n/Cirochou).













Subject A	Duration (Time)	Ratio
<i>a para</i>	 0,20 sec	1 : 1,50
<i>a: pa:ra</i>	 0,30 sec	
<i>a spa:l'i</i>	 0,16 sec	1 : 1,63
<i>a: spa:l'i</i>	 0,26 sec	
<i>æ p'æta</i>	 0,15 sec	1 : 1,46
<i>æ: p'æ:ta</i>	 0,22 sec	
<i>æ p'æna</i>	 0,16 sec	1 : 1,56
<i>æ: p'æ:na</i>	 0,25 sec	
<i>o pot</i>	 0,17 sec	1 : 1,53
<i>o: po:t</i>	 0,26 sec	
<i>e ščep'u</i>	 0,13 sec	1 : 1,39
<i>e: ščē:p'u</i>	 0,18 sec	

The length is kept best with the open vowels [a:] and [æ:] respectively. The average ratio between the short [a] and the long [a:] is 1 : 1,56. This proportion is smaller with the pair [o]—[o:] and still smaller with the pair [e]—[e:].

The average ratio between short and long vowels of the other Slovak dialects, as well as of Standard Slovak and Czech is 1 : 2 although differences occur even there.

As stated previously, the duration is the most significant feature of long Sotak vowels. But it is not the only significant characteristic. Sonagrams show that diffe-

Table II (Subject B is M. P., Papín).

Subject B	Duration (Time)	Ratio
<i>a para</i>	 0,14 sec	1 : 1,78
<i>a: pa:ra</i>	 0,25 sec	
<i>a spa:l'i</i>	 0,13 sec	1 : 1,46
<i>a: spa:l'i</i>	 0,19 sec	
<i>æ p'æta</i>	 0,15 sec	1 : 1,60
<i>æ: p'æ:ta</i>	 0,24 sec	
<i>æ p'æna</i>	 0,16 sec	1 : 1,69
<i>æ: p'æ:na</i>	 0,27 sec	
<i>o pot</i>	 0,23 sec	1 : 1,35
<i>o: po:t</i>	 0,27 sec	
<i>e ščep'u</i>	 0,11 sec	1 : 1,55
<i>e: ščē:p'u</i>	 0,17 sec	

rence between the long and short vowels is also due to the quality of sounds, due to their acoustic patterns.

X-ray pictures confirm that this is the result of the slightly different positions of articulators, especially the tongue.

This acoustic difference is smallest between the low mid vowels [a] and [a:], as shown in the sonagrams (Fig. 1 and 2). The long [a:] is more open and the second resonance bar is slightly nearer to the base line (1,7 KH) while with [a] is 2 KH.

The same difference is seen between the short and long half open front vowels [æ] (the second formant from 1,5 to 3 KH) and [æ:] (from 2 to 2,8 KH). See Fig. 3.

A greater difference can be seen in the acoustic patterns between the half closed back vowel [o] (the second bar between 1—2,4 KH) and [o:] (from 0,6 to 1 KH). The bars of the long [o:] show already a resemblance to the u-sounds. No wonder the long [o:] has changed into [o] in some Eastern Slovak dialects, or the diphthong [uo] respectively; although mostly this change has been into a short [u]. See Fig. 4 and 5.

The long [e:] is of higher frequency (the second formant from 2 to 4 KH) than the short [e] (from 2 to 3 KH), its resonance bars resembling i-sound patterns. In Eastern Slovak dialects this long vowel has changed into [i] or [e] or the diphthong [ie]. See Fig. 6 and 7.

It appears that the difference in acoustic quality between these pairs of half closed vowels [o—o:] and [e—e:] is becoming more significant in the course of time, and as this significance increases the duration of the long vowels is becoming shorter.

One or two other factors need to be pointed out. The older generation is aware that it is the longer duration that makes the difference between these pairs of sounds. The younger generation (by this we mean 30 to 60 years old) perceive the length of the

long vowels of their elders, but they themselves pronounce the sounds short. In other words:- the signal of length of the elders is perceived by the younger individuals, but it is emitted as short; as shown on the sonagrams Fig. 8 and 9, and on the Table III. The ratio between short and long vowels is 1 : 1 with the younger.

Table III (Subject C is M. G., Modra n/Cirochou).

Subject C	Duration (Time)	Ratio
<i>a para</i>	0,15 sec	1 : 1
* <i>a: pa:ra</i>	0,15 sec	
<i>æ p'æna</i>	0,18 sec	1 : 1
* <i>æ: p'æ:na</i>	0,17 sec	
<i>o pot</i>	0,17 sec	1 : 1
* <i>o: pɔ:t</i>	0,17 sec	
<i>e ščep'u</i>	0,07	1 : 1
* <i>ɛ: ščɛ:pam</i>	0,07	

* the former long vowels

It is out of scope of this paper to give all the reasons that have caused the shortening of the long Sotak vowels. We will, however, give one more.

In the Sotak dialects psycho-sociological factors have played an important part. People coming from the Sotak district into areas (especially the town Humenné) where the long vowels have been dropped and where the stress is already generally on the penultimate syllable, felt they were peculiar, or even curiosities. They were imitated and laughed at because of their speech. They therefore, consciously began to shorten their vowels. This was the reason they gave when asked them why they were trying to shorten their vowel sounds.

The children imitate the speech of the old people during the learning process, but later on they try to eliminate outstanding and handicapping variations from the norm of Eastern Slovak dialects.

And even the Standard Slovak (where the long vowels are phonemic) which they learn at school, and listen to on the radio and television does not prevent this process going on.

Liška: Length of Vowels in Sotak Dialects

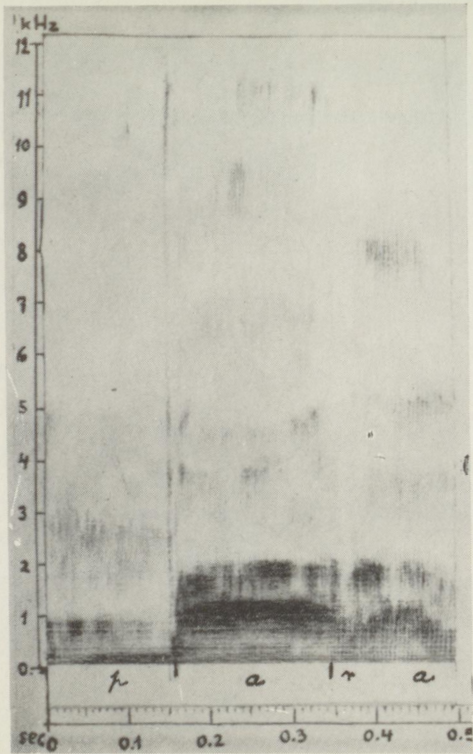


Fig. 1.

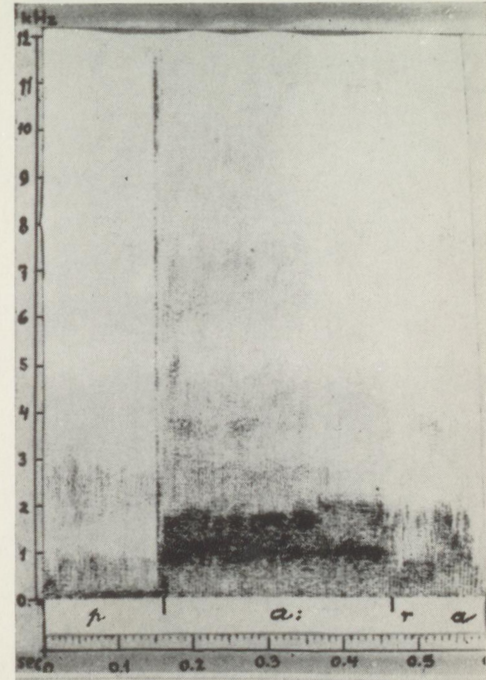


Fig. 2.

Liška: Length of Vowels in Sotak Dialects

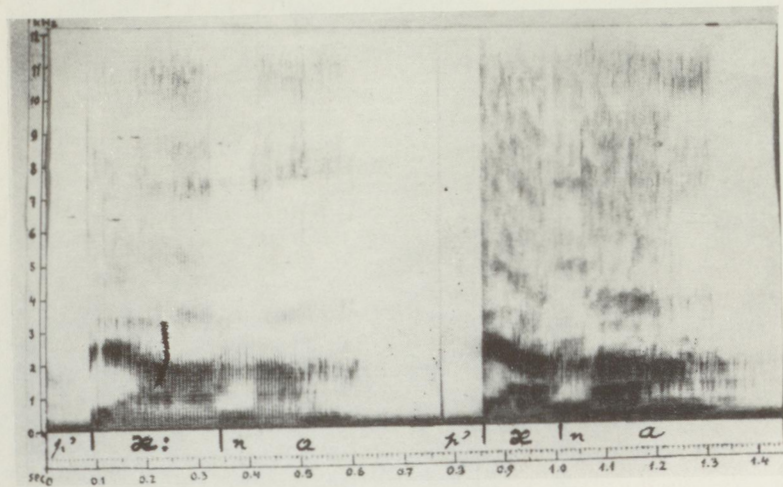


Fig. 3.

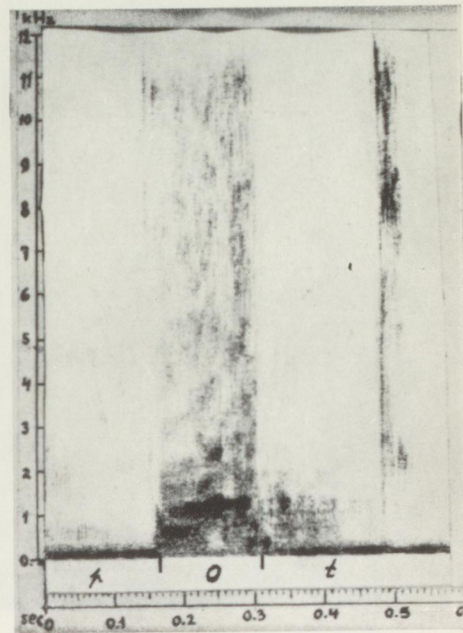


Fig. 4.

Liška: Length of Vowels in Sotak Dialects

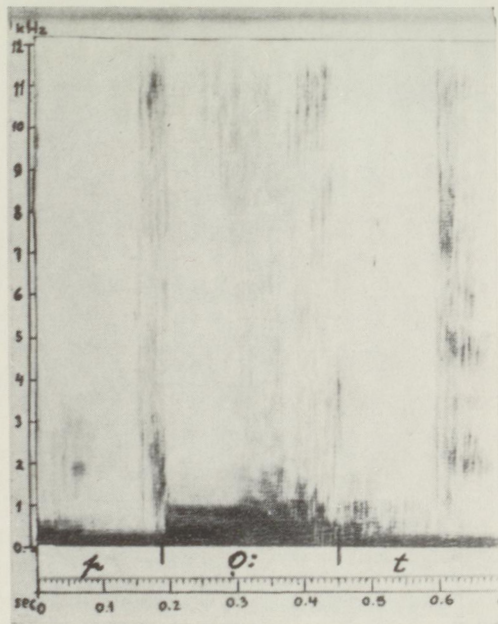


Fig. 5.

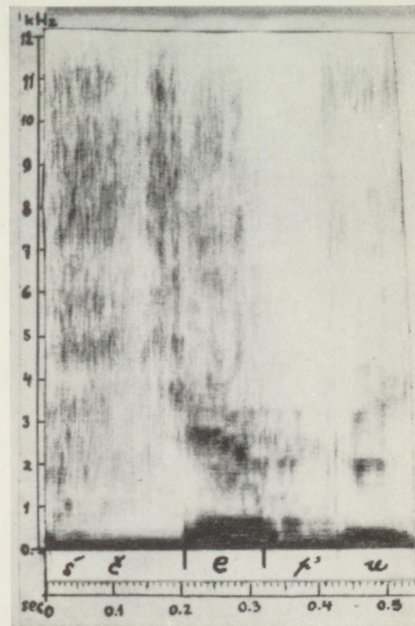


Fig. 6.

Liška: Length of Vowels in Sotak Dialects

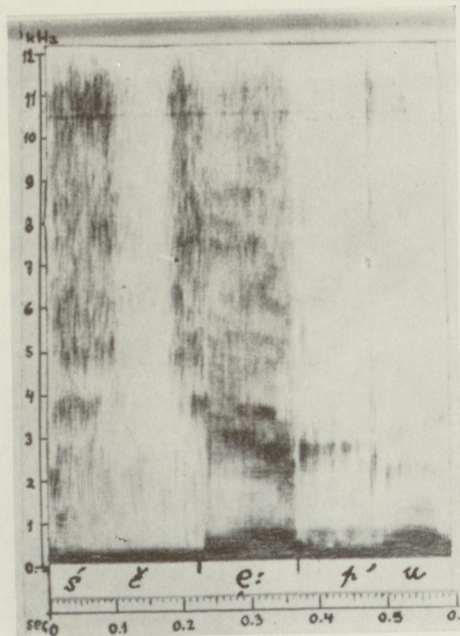


Fig. 7.

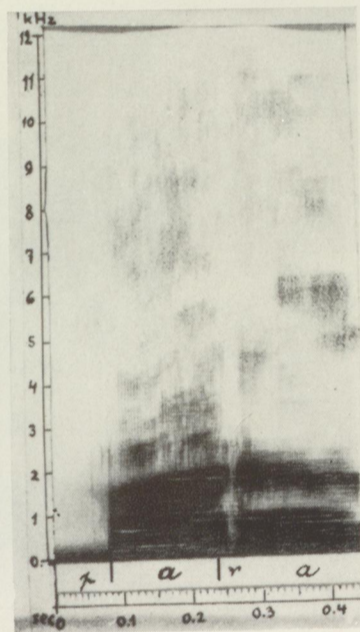


Fig. 8.

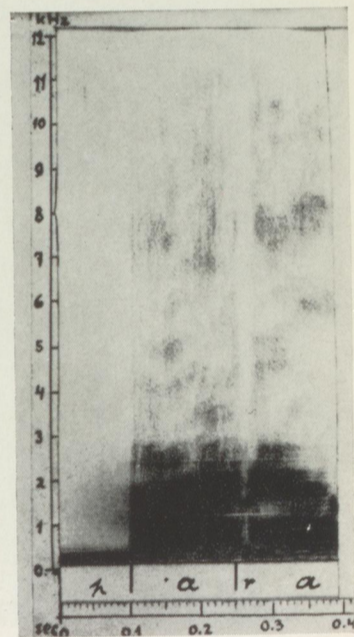


Fig. 9.

THE VOICING DIMENSION: SOME EXPERIMENTS IN COMPARATIVE PHONETICS*

LEIGH LISKER**—ARTHUR S. ABRAMSON***

Speech synthesis has made it possible to produce controlled variations along acoustic dimensions to test their perceptual relevance for phonemic distinctions. Zones of perceptual ambiguity are compared with boundaries between ranges of acoustic values measured in speech. We are interested in doing this in cross-language comparisons. For each of the languages we want to know the number of phonological categories along a dimension used in common. For languages with the same number of categories the boundaries need not be the same.

Our research has led us to believe that in many languages some phoneme categories are distinguished by the timing of glottal adjustments relative to supraglottal articulation, and that this timing relation determines not only the voicing state as narrowly defined, but the degree of aspiration and certain features associated with the so-called force of articulation as well. For word-initial stops in non-whispered speech, this relation is realized acoustically by what we have called voice onset time (VOT), i.e., the interval between the release burst and the onset of laryngeal pulsing.¹ A recent pilot study with synthetic speech demonstrated the distinctive power of VOT.² The present study takes a closer look at its perceptual relevance in three languages, English, Spanish and Thai. The resulting data also furnish a basis for discrimination experiments.

We used the Haskins Laboratories parallel resonance synthesizer, which has three

* This research was supported by the National Institute of Child Health and Human Development of the National Institutes of Health and the Information Systems Branch of the Office of Naval Research.

** University of Pennsylvania and Haskins Laboratories, New York, N. Y.

*** University of Connecticut and Haskins Laboratories, New York, N. Y.

¹ L. Lisker and A. S. Abramson, "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements," *Word* 20 (1964), 384—422; L. Lisker and A. S. Abramson, "Stop Categorization and Voice Onset Time", *Proc. 5th Intl. Cong. Phon. Sci.*, E. Zwirner and W. Bethge, eds. (Basel, 1965), pp. 389—391.

² A. S. Abramson and L. Lisker, "Voice Onset Time in Stop Consonants: Acoustic Analysis and Synthesis", *Proc. 5th Intl. Cong. on Acoustics*, D. E. Commins, ed. (Liège, 1965), Vol. Ia, Paper A 51. See also A. M. Liberman, P. C. Delattre, and F. S. Cooper, "Some Cues for the Distinction between Voiced and Voiceless Stops in Initial Position", *Language and Speech* 1 (1958), 153—166.

formant generators with variable frequencies and amplitudes, a choice of buzz or hiss excitation, or a mixture of the two, and control of the overall amplitude and fundamental frequency.

Our basic pattern was built on three steady-state formants for a vowel of the type [a]. Labial, apical and velar stops were made by adding appropriate release bursts and formant transitions to the beginning. We synthesized 37 VOT variants, ranging from 150 msec. before the release to 150 msec. after it. For voicing before the release (voicing lead), we used only low-frequency harmonics of the buzz source. For voice onset after release (voicing lag), the interval between burst and onset of pulsing was excited by hiss alone, with suppression of the first formant to simulate the effects of an open glottis. Calling the release zero time, we gave negative numbers to voicing lead and positive numbers to lag. The stimuli varied in 10 msec. steps, except for the range from -10 to +50, where we made them in 5 msec. steps. Each variant had a fundamental frequency of 114 cps, with a drop toward the end. The 37 variants were recorded in eight random orders, with two occurrences of each on each tape. The test subjects were five native speakers of Latin American Spanish, twelve of American English, and eight of Thai. Using their own orthographies, the subjects identified the stimuli with their stop phonemes.

Figs. 1-3 give the identification curves as functions of VOT values. The bars show frequency distributions of VOT values measured in speech. All the expected categories emerge, but the perceptual crossover zones do not always match very well the zones between the ranges of measured values. In Spanish (Fig. 1), (bdg) are produced with voicing lead, while (ptk) have zero VOT or short lag. The three perceptual crossovers occur to the right of these boundaries, suggesting that some other features, perhaps burst and hiss intensities or formant transitions, were not optimally set. In Fig. 2 the productions of English (bdg) show a small scattering of lead values but a concentration at zero or just after it, while (ptk) all show lag. The boundaries between ranges match the perceptual crossovers well, although there are slight discrepancies in the labials and apicals. For all three places of articulation the English perceptual crossovers have higher VOT values than the Spanish, though the differences are less than expected. Thai was chosen because for two of its places of closure it has three categories, usually called voiced, voiceless unaspirated and voiceless aspirated; the velars have only the latter two. The categories lie in the regions of voicing lead, zero VOT or short lag, and long lag respectively. The match between speech and perception (Fig. 3) is good for the left-hand boundary for the labials and apicals, showing great sensitivity to voicing lead; the right-hand boundary, however, shows the same kind of mismatch as the Spanish and English. For the velars, nevertheless, the match is perfect.

For all three languages the stop categories occupy distinct ranges along the VOT dimension. To be sure, the match between our production and labelling data is somewhat less than perfect, but this is scarcely surprising in view of the severely restricted number of variables involved in the experiment. However, despite the

likelihood that other acoustic features play a role in fixing the category boundaries studied, it seems quite clear that the timing of voice onset is a major factor in de-

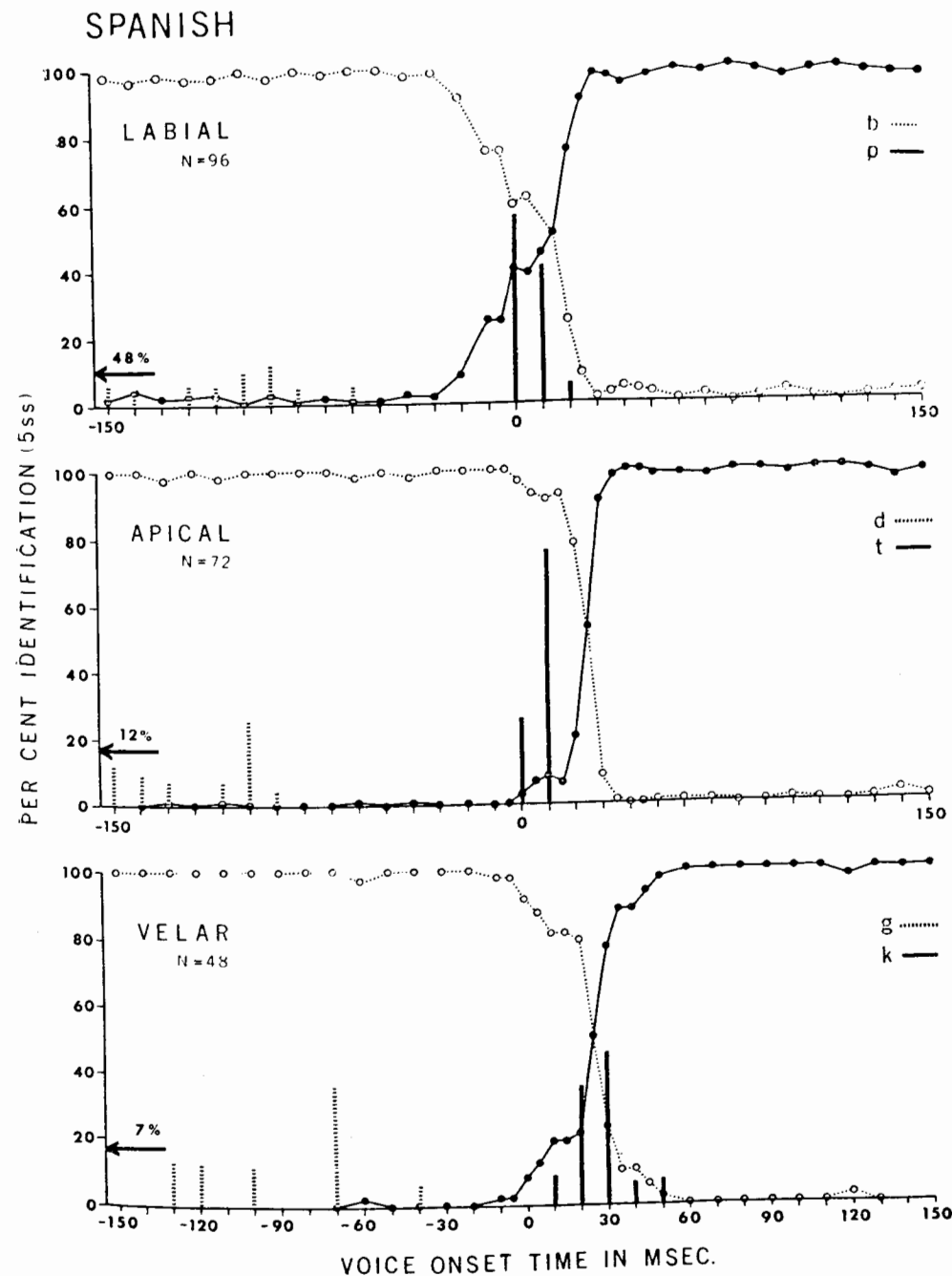


Fig. 1.

DISCRIMINABILITY ALONG THE VOICING CONTINUUM: CROSS-LANGUAGE TESTS*

ARTHUR S. ABRAMSON**—LEIGH LISKER***

In psychoacoustic experiments with speech and speech-like stimuli it has been found that listeners' ability to detect differences along a physical continuum of variation depends on their assignment of the stimuli to linguistic categories. Data on stops in synthesized stop-vowel syllables show that the stimuli are distinguished only if they are assigned to different phonemes;¹ however, such studies have been confined to English. In going into cross-language testing,² we asked two questions: (1) Is discrimination sharpened at the margins of the labelling distribution for a phoneme? (2) Is discriminability shaped by specific language experience, or does it imply general phonetic categories? Voice onset time (VOT) is of interest here because many languages use VOT, in similar but not identical ways, to distinguish either two or three stop categories.³

For this study we used a set of the stimuli prepared for the identification experiments described in the accompanying paper. Using a parallel resonance synthesizer, we made 31 syllables consisting of labial stops plus a vowel of the type [a]. VOT varied in ten msec. steps from 150 msec. (−150) before the burst, to 150 msec. (+150) after the burst. We presented these variants in triads for discrimination. In each triad two stimuli were identical and one was different. The task was to decide whether the odd one was in first, second or third position. The triads were made by pairing stimuli at two-, three- and four-step intervals along the VOT continuum, thus comparing VOT differences of 20, 30 or 40 msec. Each comparison was arranged in six permutations

* This research was supported by the National Institute of Child Health and Human Development of the National Institutes of Health and the Information Systems Branch of the Office of Naval Research.

** University of Connecticut and Haskins Laboratories, New York, N. Y.

*** University of Pennsylvania and Haskins Laboratories, New York, N. Y.

¹ A. M. Liberman, K. S. Harris, H. S. Hoffman, & B. C. Griffith, "The Discrimination of Speech Sounds within and across Phoneme Boundaries", *J. Exptl. Psych.* 54 (1957), 358—368.

² See also K. N. Stevens, A. M. Liberman & M. Studdert-Kennedy, "Cross-Language Study of Vowel Perception".

³ L. Lisker & A. S. Abramson, "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements", *Word* 20 (1964), 384—422.

(AAB, ABA, BAA, etc.) for a total of 504 triads. Six randomly ordered test series were made with only one of the six permutations in each series. The six series were split into twelve tapes of 42 triads each. Over a three-month period these tapes were played, along with identification tests, to native speakers of Latin American Spanish, American English and Thai. We shall examine the pooled data of the subjects and then remove individual differences by focusing on single subjects.

For Spanish we must report that the results are unclear. The pooled data of the six subjects do show an increase in acuity of discrimination above chance level (33.3%) in the phoneme boundary region, but there are other peaks along the continuum. The Spanish subjects often failed to discriminate between variants that they consistently distinguished in the identification tests; therefore, we believe that these results neither support nor deny the hypothesis that discrimination is more acute in the region of the phoneme boundary. Rather it seems that our Spanish speakers were not well prepared for the task. We did not communicate with them very well despite instructions written in Spanish. We hope to establish better rapport with new subjects and repeat the experiment.

The English data (Fig. 1) are clear. In the pooled data and in the graphs for two single subjects, peaks of discrimination match the phoneme boundaries established by the identification tests and indicated by the vertical lines. Discrimination is acute at all three levels of difficulty. Indeed, subject LR reaches 100% at all three levels, although the peak is flattened to a plateau at 100% only for the four-step level. The small peak for extreme values of voicing lag requires comment. Possibly this portion of the VOT dimension is near the end of the range for English /p/, bordering on non-speech sounds. We might also speculate that this small peak reflects a boundary of another type, /p/ followed by /h/ across a morpheme or word boundary. This occurs in, e.g., "Is the mop hot?" Our stimuli were not designed for such a context, but the subjects may be vaguely aware of phonetic features appropriate to a word boundary. Furthermore, there is the possible effect of the changing duration ratio between voicing lag and vowel phonation. As lag increases, the syllable duration remains fixed, but the subject's judgment of length may change enough to affect his discrimination.

The responses of the Thai group and two individuals are shown in Fig. 2. Thai has three categories of labial stops on the VOT dimension. The phoneme boundaries, shown by the vertical lines, are well-matched by the discrimination peaks. In the pooled data the two phoneme boundaries are at -20 and +40 msec. For subject UJ they are -10 and +45 msec. and for OK, -35 and +30 msec. The differences in the boundaries are largely reflected by the discrimination peaks. The small peak at extreme values of lag cited for English is also apparent here, especially for UJ. The explanations offered for English can also be advanced for Thai.

Our experiments give fairly clear answers to our questions. Although procedural problems have delayed the work on Spanish, the English and Thai data show sharpening of discrimination at the phoneme boundaries. For English at least, this might

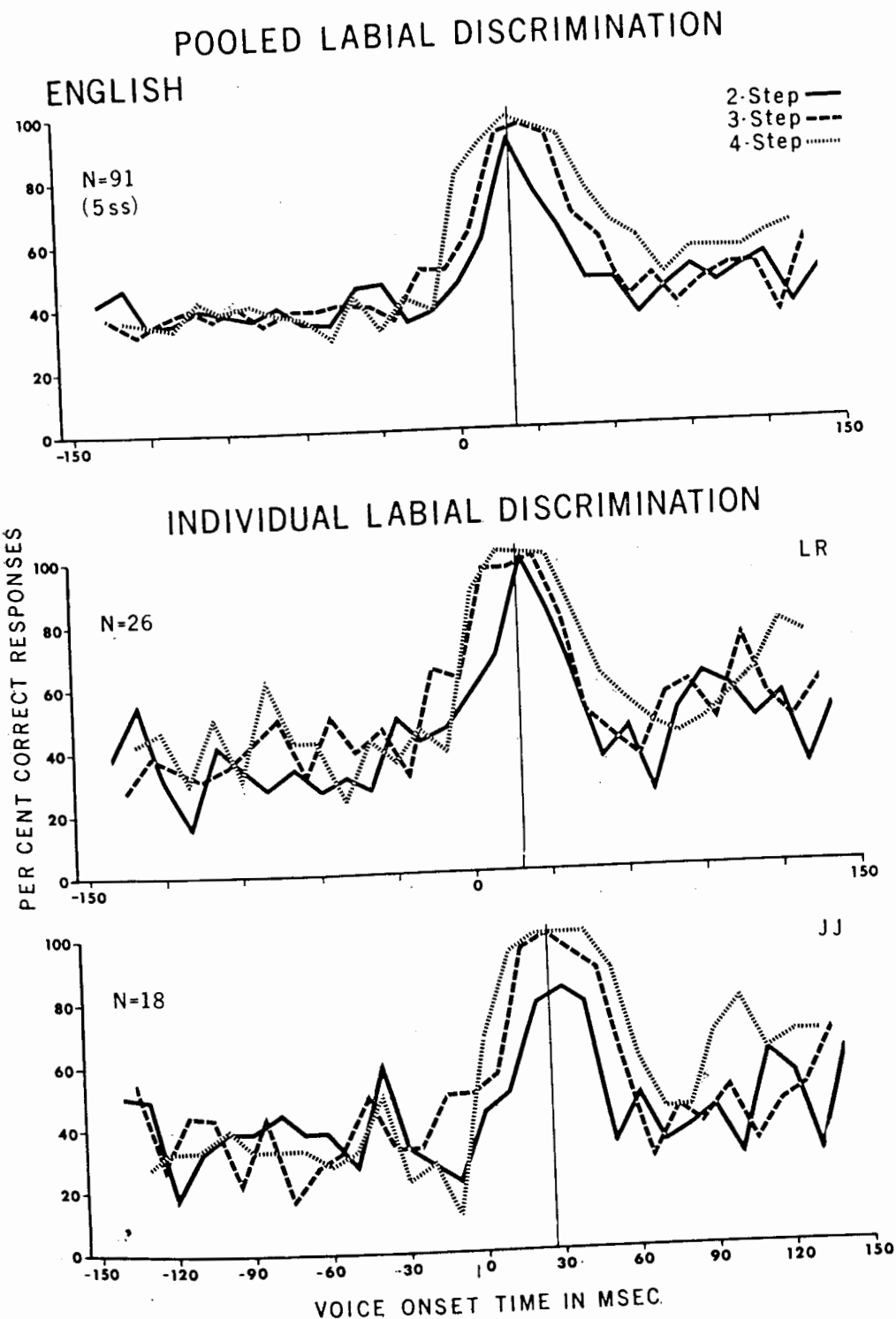
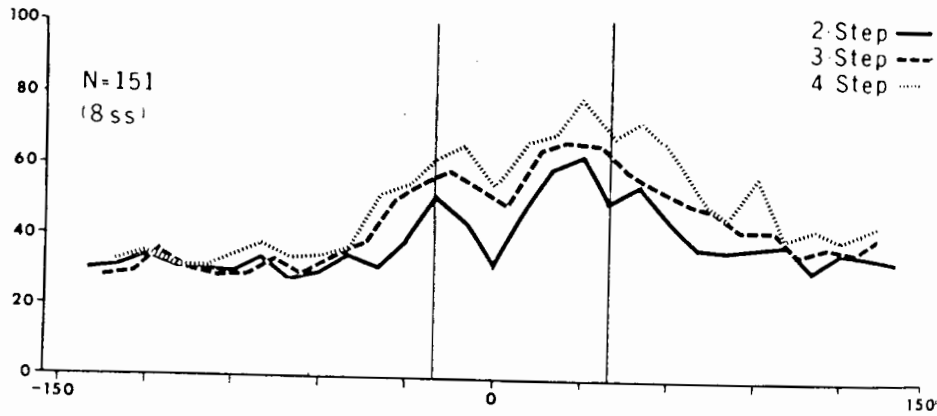


Fig. 1

POOLED LABIAL DISCRIMINATION

THAI



INDIVIDUAL LABIAL DISCRIMINATION

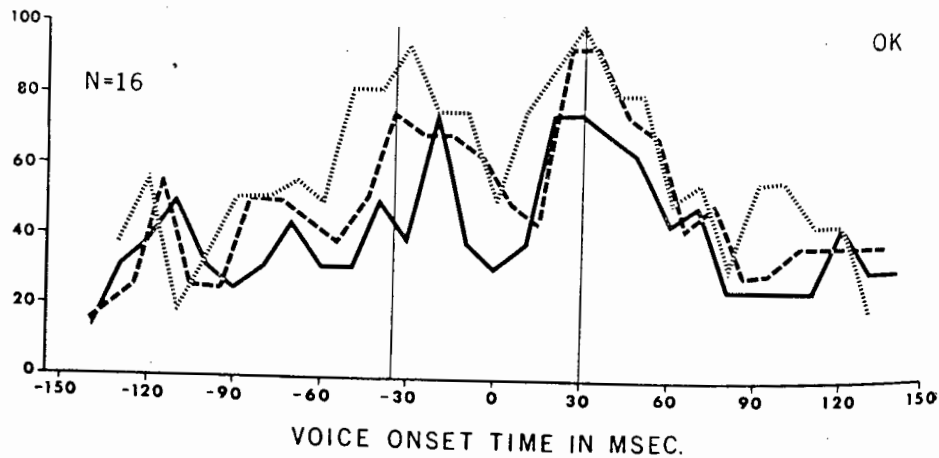
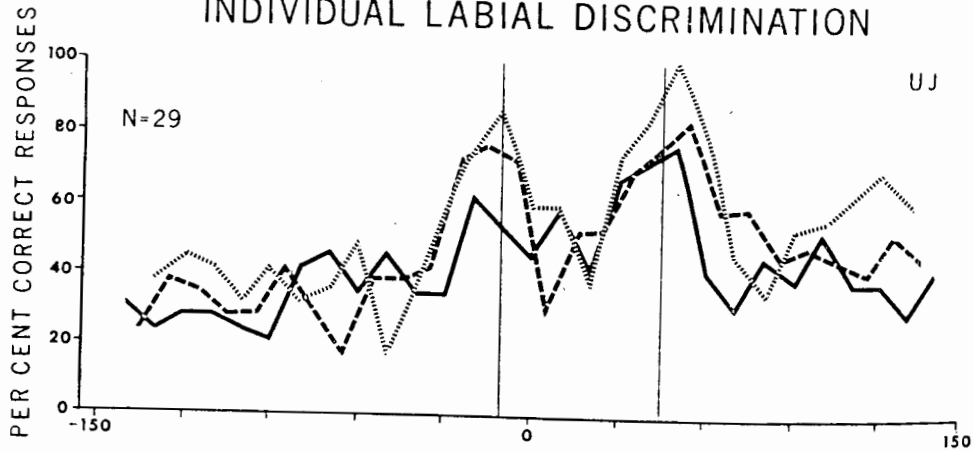


Fig. 2

have been predicted from the earlier work of Liberman *et al.*⁴ Our data suggest that discriminability is basically determined by specific language experience. There is a large peak at the boundary between the two English phonemes. The three Thai phonemes are separated by two large peaks. In both languages, for the stimuli with long voicing lag, there is a tendency to show a sharpening of discrimination close to the margin of the labelling distribution that does not abut on another phonemic category. We hope to present a report soon on similar current experiments with apical and velar stops.

DISCUSSION

Fry:

Were the Thai subjects whose individual results were given also used in collecting the analytical data? If so, was there any tendency for the data for these two speakers to be in step with their labelling results?

Abramson:

Ad Fry: No, these individual subjects were not in the United States when the analytical data were collected, nor were any of the speakers whose productions were examined for the first study still available to us at the time of the perception experiments. The possibility implied by your question cannot be ruled out until we have examined a sampling of stop productions of the present two subjects. We will try to do so soon. It would indeed be gratifying to find that the discrepancy between production and perception at one of the two boundaries is explained by individual differences in speech habits; however we doubt that this will happen.

⁴ A. M. Liberman, K. S. Harris, J. Kinney & H. Lane, "The Discrimination of Relative Onset Time of the Components of Certain Speech and Nonspeech Patterns", *J. Exptl. Psych.* 61 (1961). 379-388.

KÜNSTLERISCHE KOMPOSITION UND STIMMLICHER AUSDRUCK BEI BERTOLT BRECHT

GEERT LOTZMANN

Es soll der Versuch unternommen werden, im Sinne von Trojans ‚Lautstilistik‘¹ Ergebnisse seiner Ausdruckstheorie an Gedichten von Bertolt Brecht zu exemplifizieren. Ich wähle den ersten von Brecht zusammengestellten Auswahlband ‚Die Hauspostille‘.² 1940 schreibt Brecht: „Der Großteil der Gedichte handelt vom Untergang und die Poesie folgt der zugrundegehenden Gesellschaft auf den Grund. Die Schönheit etabliert auf Wracks, die Fetzen werden delikate. Das Erhabene wälzt sich im Staub, die Sinnlosigkeit wird als Befreierin begrüßt“.³ Wenn man die Animosität gegen alles Bürgerliche ausklammert, dann bleibt als Grundtenor Brechtscher Aussage: daß Dichtung (Lyrik) etwas sein muß, „was man ohne weiteres auf den Gebrauchswert untersuchen können muß.“³ Sein Urteil über 400 Lyriker traf den Lyrikbetrieb der Zeit zentral. „Sie entfernen sich einfach zu weit von der ursprünglichen Geste der Mitteilung eines Gedankens oder einer auch für Fremde vorteilhaften Empfindung.“³ Das Schöne und Erhabene waren nach Brecht durch die anarchisch-bürgerliche Gesellschaft korrumpiert worden. Brecht nennt das Wesen dieser Gesellschaft beim Namen.⁴

Die Lebensleere- und -lüge beantwortet er mit einem „rauschhaft-animalischen Daseins- und Naturgefühl“⁵ Baalscher exzessiver Vulgarität oder er verfremdet das kreatürliche Leiden und Absterben zum „Grunderlebnis des sich Treibenlassens“⁶ fern jeder romantischen Verklärung, lediglich auf den Verstand und die Nutzenwendung gerichtet.

Die ‚Hauspostille‘ spielt nicht bloß mit Reminiszenzen aus dem Erbauungsbuch, sondern ist neben anderem auch umfunktionierte „Theologie“.⁷ Brecht führt uns 400

¹ Trojan Felix: Die Ausbildung der Sprechstimme, Wien 1948; Der Ausdruck der Sprechstimme, Wien 1952; Sprachrhythmus und vegetatives Nervensystem, Wien 1951; Der Ausdruck der Sprechstimme im Deutschen, Frankfurt 1954; Die Kunst der Rezitation, Wien 1954; Psychodiagnostik der Persönlichkeit an Hand von Sprachwerken, in: *Sprachforum* 2/1956, H. 2.

² Brecht, Bertolt: *Bertolt Brechts Hauspostille*, Berlin 1927.

³ Brecht, Bertolt: *Über Lyrik*, edition suhrkamp 70, Frankfurt 1964.

⁴ Benjamin, Walter: *Versuche über Brecht*, edition suhrkamp 172, Frankfurt 1966.

⁵ Grimm, Reinhold: Bertolt Brecht, in: *Sammlung Metzler*, Stuttgart 1961.

⁶ Esslin Martin: *Brecht. Das Paradox des politischen Dichters*, Frankfurt 1962.

Jahre nach dem Erscheinen von Luthers Kirchen- und Hauspostille von 1527, fünf Lektionen zum Gebrauch im gegenwärtigen Alltag vor: Bittgänge, Exerzitien, Chroniken, Mahagonnygesänge, Die Kleinen Tagzeiten der Abgestorbenen. Es geht ihm um die Verneinung vorgegebener Inhalte und die Formulierung von Gegenthesen mit der Absicht, die alten geistigen Inhalte der sakralen Erbauungsbücher zu zerstören. Die 1. Lektion besteht aus 9 ‚Bittgängen‘, die sich an „das Gefühl des Lesers“⁸ wenden. Indem die Bitten oder Gebete auf die Existenzweise der bürgerlichen Gesellschaft übertragen werden, erreicht der Dichter mittels der Umfunktionierung die Säkularisierung einer ursprünglich christlichen Zeremonie. Nicht der Himmel allein kann Abhilfe schaffen, sondern auch der Mitmensch ist aufgefordert zu ändern: ‚Doch ihr, ich bitt euch, wollt nicht in Zorn verfallen / Denn alle Kreatur braucht Hilf von allen.‘ (Von der Kindesmörderin Marie Farrar)

Das ‚Akuem‘ Bitte gehört nach Trojan zur trophotropen Funktionsrichtung mit dem lautstilistischen Sprechtakt der Ruhehaltung und der Gerichtetheit. Wenn man die Bitte noch durch die Akueme Mitleid (Das Schiff), Güte — die ich hier im Sinne von Abbitte meine — (Morgendliche Rede an den Baum Green), körperliches Wohlbehagen verbunden mit der sinnlichen Lust — allerdings als ungerichtetes Akuem — (Vom François Villon) ergänzt, so stößt man beim Sprechen dieser Gedichte auf die Merkmale der von Trojan verschiedenenorts beschriebenen A-Klasse,⁹ die sich bei Einsatz der Schonstimme vor allem durch die faukale Weite, ruhig fließenden Atem, weiche Stimmeinsätze, geringe Artikulationsspannung, den Legato-Charakter und eine Assimilationstendenz bestimmen. Weitere Merkmale sind vorherrschende Innenspannung, dunklere Klangfärbung, Betonung des Melos und relativ ruhiges Zeitmaß.¹⁰ Natürlich sind die Signale bei den einzelnen ‚Bittgängen‘ mehr oder weniger stark ausgeprägt bzw. ist ihr konstitutiver Charakter Abweichungen unterworfen. In der 2. Lektion vereinigt Brecht unter den ‚Exerzitien‘ 14 Gedichte mit umkehrendem Anleitungs-, Unterweisungs- und Belehrungscharakter, d. h. die christliche Katharsis-Prozedur verwandelt sich in eine „vulgärmaterialistische Übung im Lebensgenuß“.¹¹ Die Diktion verrät gleichzeitig die „Anwendung strenger gesellschaftskritischer Maßstäbe“,¹² so daß sich die Gebrauchsanleitung mehr an den

⁷ Mayer, Hans: *Bertolt Brecht und die Tradition*, Pfullingen 1964.

⁸ *Bertolt Brechts Hauspostille*, Anleitung zum Gebrauch der einzelnen Lektionen.

⁹ Trojan, Felix: Die Ausdruckstheorie der Sprechstimme, in: *Phonetica* 4/1959; Sprech-erziehung auf entwicklungsgeschichtlicher Grundlage, in: *Z. f. Phonetik u. Allg. Sprachw.* 12/1959; Zur Kategorienlehre der Sprechkunde, in: *Sprechk. u. Sprecherz.* Bd. IV, Emsdetten 1959; Stimmlicher Ausdruck und Sprache, in: *Proceedings... 4. Congr.*, Helsinki 1961; derselbe mit Winckel, Fritz: Elektroakustische Untersuchungen zur Ausdruckstheorie der Sprechstimme, in: *Fol. Phoniatria* 9/1957.

¹⁰ Stelzig, Helmut: Sprechwissenschaftliche Analysen der Klangstruktur zur Lyrik der Gegenwart, in: *Wiss. Z. d. Univ. Halle*, Ges.-Sprachw. Reihe, IX/1960.

¹¹ Schuhmann, Klaus: Der Lyriker Bertolt Brecht. 1913—1933, in: *Neue Beiträge z. Literaturwissenschaft*, Bd. 20, Berlin 1964.

Verstand wenden muß. „Es ist vorteilhaft, ihre Lektüre langsam und wiederholt, niemals ohne Einfalt, vorzunehmen“⁸. Aus den Hinweisen soll mancher Aufschluß über das Leben zu gewinnen sein: Der Branntweinhändler gibt sein bisher sinnlos verpraßtes Geld den Waisenkindern (Vorbildliche Bekehrung eines Branntweinhändlers); ‚Von der Freundlichkeit der Welt‘ belehrt uns über den unausweichbaren Lebenslauf bis zum Tode; Anweisungen zur passiven Lebenshaltung sind die Exerzitien ‚Vom Klettern in Bäumen‘ und ‚Vom Schwimmen in Seen und Flüssen‘. Das Fazit dieser ‚geistlichen Übungen‘ kulminiert im ‚Großen Dankchoral‘, einer Lobpreisung von Nacht und Finsternis, Baum und Aas, des schlechten Gedächtnisses des Himmels und schließlich der Kälte und des Verderbens. Dieser Gruppe von Gedichten wäre die ideotrope Einstellung zur Umwelt zuzuordnen, die bestimmt ist durch Hemmungen des Affektiven und das Hervortreten des Logischen, Verstandesmäßigen (Klasse C). Mäßige Gespanntheit ist verbunden mit faukaler Weite bei auffälliger Personen- bzw. Sachgerichtetheit. „Im Stimmausdruck und der Sprechgestaltung treten die emotionalen Faktoren zugunsten der rationalen zurück: Die Klangstruktur wirkt durch stärkere Einordnung der einzelnen Klangfaktoren und die Einschränkung der Ausdruckspolarität ausgeglichener, sachlicher“.¹³ Der Brechtsche Gestus, der sich in Körperhaltung, Gesichtsausdruck und Tonfall ausdrückt, unterstreicht die Gerichtetheit, denn alle Äußerungen zielen auf den Mitmenschen. „Weniger soll die Gestalt eines Satzes... geprägt sein von dem Inhalt, den sie transportiert, als von Haltung, Gefühlen und Verfassung der Person, die spricht und ihrer Beziehung zu dem Inhalt.“¹⁴

Als dritte Gruppe führt Brecht ‚Die Chroniken‘ in die Hauspostille ein. Indem die Worthülse aus dem Alten Testament übernommen wird, erfahren wir die Stories der Outcasts in Balladenform mit Moritaten- und Bänkelsangunterton. Nicht mehr Könige sind die Helden, sondern die durch die Gesellschaft Entrechteten mit der Gebärde der vital-anarchischen Lebensbewältigung. Die ‚Abenteurer‘, die die Jugend, aber nicht ihre Träume vergessen haben; ‚die Männer von Fort Donald‘, die Schienen legten, bis sie eroffen; die ‚Seeräuber‘, die noch im Sterben um ihr einzig seligmachendes Element flehen; der unmenschliche Lebensgang der ‚Hanna Cash‘, die nur im Himmel Gerechtigkeit finden kann; der zu Tode gerittene Kosaken-Hetman ‚Mazepa‘ und schließlich ‚der Soldat mit dem Messer im Gurt‘, der, nicht hörend auf den Rat des weisen Weibes, vergeht wie der Rauch. Nach Brecht ist diese Lektion „in den Zeiten der rohen Naturgewalten, ... der Gefahr ... und beispiellosen Verfolgung“ zu lesen.⁸ Diese ironischen Lesevorschriften sind ernst gemeint. Fern jeder Gefühlslyrik erlebt sich Brecht in der Maske des „Erbauungspoeten als Gesellschaftssatiriker“.¹⁵ Trotz einer Grundstimmung, die mit Umschreibungen wie Nicht-

¹² Jens, Walter: *Statt einer Literaturgeschichte*, Pfullingen 1957.

¹³ Stelzig, Helmut: a. a. O., 49.

¹⁴ Klotz, Volker: *Bertolt Brecht. Versuch über das Werk*, Bad Homburg v. d. H. 1957.

¹⁵ Heselhaus, Klemens: *Deutsche Lyrik der Moderne*, Düsseldorf 1962.

mehr-können, Vergessen, Vergehen, Hinsinken, Fallen, Untergehen, Sterben, Ewige-Ruhe-haben zu kennzeichnen ist, haben doch gerade die ‚Chroniken‘ durch die Umstrukturierung traditioneller Lyrismen — die ‚Erinnerung an die Marie A‘ scheidet hier aus — „in einer grausamen Gegenwart“ und, für den Dichter, „unbarmherzigen Zukunft“¹⁶ viel von jener Tendenz, die durch ihre brutale Konfrontierung mit gegenwärtigen Mißständen die positive Veränderung erstrebt. Die umfunktionierte Erbauungstheologie steht im „Rahmen einer umfassenderen Gesellschaftskritik, nicht Blasphemie.“¹⁷

Die ergotrope Funktionseinstellung, die man allgemein als aktives Eingreifen in ein Geschehen, als Werkgerichtetheit mit dem Grundtypus der Kraftstimme charakterisieren kann (B-Klasse), entspricht am nächsten dieser Gedichtgruppe. An Einzel-Akuemen wären allgemeines körperliches Mißbehagen, körperlicher und seelischer Schmerz, aber auch Verachtung und Zorn zuzuordnen. Der Stimmausdruck äußert sich prinzipiell in einer extravertierten Sprechspannung und damit der Tendenz zu dynamischer Akzentuierung. Die Klangfarben sind heller, das Redetempo ist schneller. Zu beobachten ist die Neigung zu festen Stimmeinsätzen und stakkatierter Rhythmisierung; der Konsonantismus ist ebenso dominierend wie die Rachenenge.

Es ist notwendig, auch im Sinne von Höffes und Fährmanns Vorgehen¹⁷ das hier Angedeutete durch Untersuchungen der einzelnen Gedichte über das Deskriptive hinaus zu verfolgen.

¹⁶ Schöne, Albrecht: Bertolt Brecht. Erinnerung an die Marie A, in: *Die deutsche Lyrik II*, Düsseldorf 1959.

¹⁷ Höffe, Wilhelm L.: Sprachlicher Ausdrucksgehalt und akustische Struktur, in: *Wiss. Z. d. Univ. Jena, Ges.-Sprachw. Reihe: 6/1956/57*, H. 6: Buchausgabe: Ratingen 1966 und Fährmann, Rudolf: *Die Deutung des Sprechausdrucks*, Bonn 1960.

DIE ALPHABETSCHRIFT UND DAS PROBLEM DER LAUTSEGMENTIERUNG

HELMUT LÜDTKE

I. Die klassische Phonemtheorie beruht — ebenso wie die Lehre der Junggrammatiker — auf folgender (ausgesprochener oder unausgesprochener) Voraussetzung: der kontinuierliche Redestrom (*chaîne parlée*) gliedert sich in eine Abfolge kleinster Zeitsegmente (die von den Junggrammatikern als „Laute“, von den Strukturalisten als „Phoneme“ bezeichnet werden); diese Gliederung in kleinste Zeitsegmente, auf der wissenschaftliche Lautschriften und historische Alphabete beruhen, sei natürlich und universal; daß die Segmentierung so und nicht anders erfolgt, müsse in der Natur des Redestromes begründet liegen.

Gegen die soeben skizzierte Auffassung, die man wohl als die herrschende Lehre bezeichnen darf, sind in den letzten Jahren mehrere schwerwiegende Argumente vorgebracht worden. Zunächst erwuchs bei der Behandlung des Problems der phonematischen Wertung mancher Lautkomplexe (Diphthonge, Affrikaten) die Erkenntnis, daß hier eine eindeutige Lösung prinzipiell unmöglich ist. Sodann ergaben Forschungen zur Sprachsynthese, daß man zwar Phonemsegmente künstlich erzeugen kann, ihre Zusammensetzung jedoch keine verständliche Rede ergibt. Bei einem anderen Segmentierungsverfahren hingegen, das nicht auf der klassischen Phonemtheorie beruht, läßt sich verständliche Rede synthetisch erzeugen. Schließlich haben die Versuche von Truby zu der Schlußfolgerung geführt, daß dem Phonembegriff überhaupt kein physikalisches Korrelat (weder ein artikulatorisches noch ein akustisches noch ein auditives) direkt entspricht.

Wenn die eben genannten experimentellen Daten stimmen, dann muß man sich fragen, wie unsere Wissenschaft überhaupt zu dem Begriff des Sprachlautes oder Phonemsegmentes gekommen ist. Der Schluß liegt nahe, daß die jahrtausendelange Gewohnheit der Alphabetschrift den ersten Phonetikern diesen Begriff einfach suggeriert hat. Dem steht freilich die landläufige Ansicht entgegen, wonach die Gliederung des Redestromes primär, die Alphabetschrift sekundär sei. Es wird nicht leicht sein, in dieser Kontroverse einen strikten Beweis zu führen; jedoch soviel ist klar: die Phonemtheorie verliert ihre Stütze, wenn es gelingt, die Erfindung des Alphabets anders zu motivieren als durch die Annahme einer vorgegebenen natürlichen Segmentierung des Redestromes. Das soll im folgenden versucht werden.

II. Zunächst einige Vorbemerkungen zur Geschichte der Schrift. Es lassen sich

— grob gesehen — vier verschiedene Prinzipien der Verschriftung von Sprache unterscheiden:

1. Piktographie = Wiedergabe von Sachverhalten durch motivierte Zeichen (das ist erst eine Vorstufe der eigentlichen Schrift)
2. Logographie = Wiedergabe von Monemen durch arbiträre Zeichen
3. Syllabographie = Wiedergabe von Silben durch arbiträre Zeichen
4. Alphabet = Wiedergabe von Teilmengen des Schallkorrelates von Monemen (durch arbiträre Zeichen)

Logographie und Syllabographie sind *natürliche* Verschriftungsprinzipien, d. h. sie beruhen auf universalen sprachlichen Vorgegebenheiten, und zwar die Logographie auf dem Monem als der kleinsten bedeutungstragenden Einheit, die Syllabographie auf der Silbe als universalem Schallkomplex. Die natürlichen Verschriftungsprinzipien sind a priori polygenetisch. Tatsächlich sind sowohl logographische als auch syllabographische Systeme an verschiedenen Stellen der Erde zu verschiedenen Zeiten geschaffen worden, ohne daß sich Einflüsse nachweisen lassen. Demgegenüber ist das Alphabet (als Prinzip) offensichtlich nur ein einziges Mal erfunden worden, und zwar im Gebiet der Ostküste des Mittelmeeres (im Raum Syrien-Libanon-Palästina-Sinai) zwischen 1500 und 1200 v. Chr.

Der mutmaßliche Erfinder des Alphabets war Semit. Um die Methode seines Vorgehens nachvollziehen zu können, müssen wir uns einige Grundgegebenheiten des semitischen Sprachbaues vor Augen führen. Das geschieht am besten anhand eines Beispiels. Nehmen wir also eine lexikalische Familie, die wir durch die Radikale Š L M bzw. (je nach Sprache) S L M symbolisieren und die in den verschiedensten semitischen Idiomen reichhaltig vertreten ist. Im Ugaritischen — um die älteste alphabetisch verschriftete Sprache zu nennen — finden wir sie mit folgenden Bedeutungen:

Š L M	unversehrt sein
	Friede
	Friedensgabe
	Opfergabe
	ausgezahlt, beglichen
	Gott der Abenddämmerung

In der heute am weitesten verbreiteten semitischen Sprache, dem Arabischen, finden wir:

S L M	
salima	unversehrt sein
salm	} Friede
silm	
salām	Heil
salāma	Tadellosigkeit

satīm	} sicher
sālim	
Sulaimān	Salomo
islām	Hingabe an Gott

Vor der Erfindung des Alphabets war die kleinste Einheit, nach der man den Redestrom gliederte, entweder das Monem oder die Silbe. Die beiden daraus resultierenden Schriftprinzipien, Logographie und Syllabographie, sind unter sich heterogen, unterscheiden sich aber gemeinsam dadurch vom Alphabet, daß sie eine kleinere Zeichenmenge im Text, eine größere im Repertoire erfordern; auf diese Weise sind sie weniger ökonomisch. Der geringere Lernaufwand des Alphabets ist der Grund für dessen Erfindung oder doch zumindest für seine Durchsetzung und Weiterverbreitung.

Das Problem, dem sich der Alphabetfinder konfrontiert sah, lautete: eines der Elemente des Redestromes, entweder Monem oder Silbe, in noch kleinere Einheiten zerlegen. Bei der Lösung kam ihm der semitische Sprachbau entscheidend zu Hilfe. Aus den vorhin angegebenen ugaritischen und arabischen Wörtern läßt sich unschwer ein allen gemeinsames Monem isolieren, das wir *X* nennen wollen, und ferner eine Reihe *verschiedener* Moneme, die wir als Y_1, Y_2, Y_3 usw. bezeichnen. Dann ist jedes der genannten Wörter eine Summe $X + Y_1$ bzw. $X + Y_2$ bzw. $X + Y_3$ usw.

Bemerkenswert am semitischen Sprachbau ist, daß jene Summen aus *X* und *Y* sich nicht als zeitliche Abfolgen *YX* oder *YX* auffassen lassen (wie etwa lat. *amicus* oder frz. *l'ami*), sondern daß *X* und *Y* ineinander verzahnt sind.

In dieser semitischen Eigenart lag die große Chance für den Alphabetfinder. Ein Komplex wie [*salima*] gliedert sich nicht nur in ein Semantem *X* und ein Morphem *Y*, sondern darüberhinaus zerfallen *X* sowohl wie *Y* infolge der Verzahnung in die Teilmengen $X' - X'' - X'''$ bzw. $Y' - Y'' - Y'''$, wobei der gesamte Komplex [*salima*] als Abfolge $\frac{X' X'' X'''}{Y' Y'' Y'''}$ aufgefaßt werden kann. Durch den semitischen Sprach-

bau wird also eine Zerlegung des Monems in Teilmengen mit festliegender Abfolge suggeriert. Aus dieser Gegebenheit entstand — in Verbindung mit der Erkenntnis, daß *X* ungleich informationshaltiger ist als *Y* und daß die natürliche Sprache ein hohes Maß an Redundanz enthält — ein Verschriftungsprinzip, welches jeder durch den semitischen Sprachbau gegebenen Teilmenge von *X* je ein arbiträres Zeichen zuordnet.

Damit war zum ersten Mal eine Analyse des Monems in noch kleinere Einheiten gelungen. Freilich sind die so gewonnenen Einheiten, von den Griechen *γράμματα*, von den Römern *litterae*, von den Junggrammatikern *Larve* und von den Strukturalisten *Phoneme* genannt, spezifisch semitisch und nicht universal. Daß die Alphabetschrift sich dennoch auf der ganzen Welt durchsetzt, liegt einerseits an der Ökonomie dieses Systems in Vergleich zu den anderen Verschriftungsprinzipien, andererseits am Scharfsinn jenes griechischen Umschöpfers, der das Alphabetsystem auch für nicht-semitische Sprachen verwendbar machte.

Grotesk erscheint, daß Generationen von Phonetikern durch die alphabetische Tradition genasführt und auf eine falsche Fährte gelenkt worden sind.

DISCUSSION

M. Cohen:

Comme l'a remarqué Antoine Meillet les hommes qui ont inventé et perfectionné l'écriture ont fait oeuvre de vrais linguistes en passant de l'analyse du discours en mots à la reconnaissance finale du phonème dans l'écriture alphabétique. Mais il s'est posé partout des questions orthographiques dont l'histoire est très compliquée.

On peut citer comme exemple l'écriture éthiopienne à syllabisme secondaire, où les lettrés comptent la longueur des mots par le nombre de caractères syllabiques, sans distinguer les racines. Pour la reconnaissance des mots complets on emploie ou on n'emploie pas des signes de séparation.

On trouve le point ou la barre de séparation dans les inscriptions phéniciennes et sudarabiques et dans les inscriptions latines. Les manuscrits grecs et latins ont employé l'écriture continue. Ce n'est qu'au XI^e siècle que s'est stabilisée l'habitude de mettre des intervalles. Dès lors nous avons pris l'habitude d'une double reconnaissance, plus ou moins parfaite, des sons par les lettrés et d'autre part des mots (ou idées) par les blancs.

Martinet:

Tous les systèmes sont en mouvements. Par application des règles existantes on n'arrive pas à se décider pour un ou deux phonèmes, il y a possibilité de deux solutions, certains traits distinctifs étant dans un cas attribués à deux segments successifs. Sur la question de l'origine de l'alphabet il faut noter que les syllabaires préexistants détruisaient l'unité graphique du mot sémitique; la non-notation des voyelles („alphabet“ phénicien) a permis de restituer cette unité. Le passage au grec a imposé la notation des voyelles.

Kohler:

Ist das Devanagari eine Silbenschrift? Wenn ja, dann ergeben sich Schwierigkeiten in Ihrer Argumentation. Das D. enthält nämlich Kriterien, die man als alphabetisch bezeichnen muß. Zeichen weisen auf Konsonant + Vokal hin, und andere Vokale werden durch Modifikationen dieser Zeichen in konsistenter Weise angedeutet (während andere Konsonanten neue Zeichen erfordern).

Meinecker:

Daß der Einfluß der Alphabetschrift auf die linguistische Forschung als ein Hindernis bei der Erkenntnis der Lautstruktur so klar herausgestellt worden ist, kann nur begrüßt werden. Jedoch würde ich die entgegengesetzte Schlußfolgerung aus den vorgetragenen Tatsachen ziehen: Alle Schwierigkeiten bei der Phonemisierung beweisen nicht, daß es keine Phoneme gibt, sondern daß unsere „klassische“ Phonemisierung nicht immer die sprachlich relevanten Faktoren richtig erfaßt. Die kleinsten Einheiten der Phonemisierung dürfen nicht so weit abstrahiert werden, daß dadurch wesentliche Merkmale in der Struktur der Sprache verloren gehen.

Richter:

Eine Kritik an der Universalität des Phonems dürfte eher in semiotischer als in historisch-geographischer Hinsicht berechtigt sein. Es ist fraglich, ob das Phonem eine geeignete Einheit

für die Segmentierung des Sprachsignals (Auteme nach Tillmann) oder seiner Wahrnehmung ist. Der Diskussionsredner glaubt, daß das Phonem auch bei Zugrundelegung einer äußerst abstrakten Konzeption Prinzipien der Zeichendifferenzierung und der Segmentierung entspricht, die für die motorisch-artikulatorische Seite der lautsprachlichen Kommunikation spezifisch sind.

Vachek:

It may be that in some languages the phonemic frame of reference does not serve so well as some other frame (cf. the arguments of J. R. Firth's school). Yet the unsolved problems pointed out by Prof. Lüdtke may be due to the well-known fact of the phonemic system being in motion—see especially the problem of the mono or biphonemic value of diphthongs. The big spread of the graphemic system of writing is not due to facts of economy alone but also to the presuppositions for its spread found in the very character of language systems.

Lüdtke:

ad Kohler: Das Devanagari ist eine Pseudo-Silbenschrift, d. h. etwas, das sich mit dem einfachen Begriffspaar „Alphabet/Silbenschrift“ nicht genau erfassen läßt. Schriftgeschichtlich gesehen, setzt das Devanagari die vorangegangene Erfindung des (semitischen) Alphabets voraus; es ist ja auch chronologisch jünger. Man kann das Devanagari (und alle typologisch verwandten Schriftarten) als sekundäre Umgestaltung des Alphabets betrachten.

АУДИРОВАНИЕ РЕЧЕВЫХ СООБЩЕНИЙ С ТОЧКИ ЗРЕНИЯ ИХ СТРУКТУРЫ И РЕЛЕВАНТНОСТИ ДЛЯ ОПЕРАТОРА

И. М. ЛУЩИХИНА*

Изучение речевого потока, идущего по нескольким каналам связи, входит в проблему многоканальной речевой коммуникации, представляя собой специальный раздел инженерно-психологического изучения различных видов сигнализации. Любые методы анализа речевых сообщений в этом случае должны сочетаться с методами анализа задачи, стоящей перед оператором.

Поскольку задачей нашего экспериментального исследования было выяснение закономерностей выборного, или селективного аудирования, мы обратились к гипотезе о фильтрах Бродбента. Бродбент полагает, что на входе нервной системы человека есть своеобразный фильтр, который не пропускает информацию, превышающую объем канала (в данном случае — объем человеческого мозга). Очевидно, первый, наиболее элементарный этап в сложном процессе опознания нужного сообщения — это более простая проблема различения физической окраски входного канала (его тембра, интенсивности и т. д.). Следующий этап — идентификация слов и значений по отношению к выбранным сигналам — представляет собой гораздо более сложную задачу, входящую в область семантических исследований.

Целью работы, проводимой в нашей Лаборатории, явилось изучение соотношения этих двух ступеней в аудировании релевантных речевых сообщений и более детальный анализ тех затруднений, которые возникают у слушателя при требовании дифференциации по смыслу. Исследование проводилось на материале фраз диспетчеров аэропорта, поступающих с различной интенсивностью — от 60 до 80 дб — к одному принимающему оператору. Была использована методика артикуляционных испытаний.

Результаты эксперимента подтвердили априорное предположение о том, что дифференциация по смыслу является гораздо более сложной задачей для испытуемого, чем дифференциация по одним акустическим характерис-

* Ленинградский государственный университет. Лаборатория инженерной психологии.

тикам, в данном конкретном случае — по изменениям интенсивности фраз. Так, если в первой серии общее смысловое восстановление фраз составляло 59 %, то во второй серии оно возросло до 83 % при уровне значимости в 0.01.

Поскольку дифференциация по смыслу является сложным вопросом входящим в область семантической информации, мы ограничиваемся синтаксическими рассуждениями, приводящими к рассмотрению отдельных элементов синтаксических структур.

В проведенных нами психолингвистических экспериментах подтвердилось положение о наиболее значимых элементах сообщения, составляющих его остоу. Проведенные подсчеты позволяют заключить, что наиболее существенными для носителя языка оказались предикативные связи, комплетивные связи типа управления (т. е. прямо или косвенно направлены на объект действия) и координационные связи (типа связей между собой однородных членов).

Обратившись далее к анализу грамматических классов, мы составили, пользуясь приемами дистрибутивного анализа, таблицу вероятности появления последующего класса за фиксированным исходным. Таблица, составленная на материале специальных текстов, дала первое, упрощенное представление о размерах структурных разрывов между различными членами выражения. Оказалось, что наибольшая вариативность, а, следовательно, и наибольшие структурные разрывы возможны в следующих случаях: после существительного, глагола, после терминальной и внутренней пунктуации и после числительного.

Эти рассуждения привели к мысли о том, что грамматическая система для информанта является сложным стохастическим процессом. Чем более вероятным будет появление той или иной грамматической категории, тем больше оснований рассчитывать на правильную реакцию слушателя. Наибольшая же вариативность вызовет наибольшие затруднения у оператора, который сталкивается в таком случае с проблемой множественного выбора.

Затруднения слушателя мы определяли с помощью пауз нерешительности, которые, врываясь в нормальный речевой поток, способны не только замедлить его, но и остановить.

Распределение пауз нерешительности в ответах испытуемых ставилось в зависимость только от одного фактора — характера грамматической категории. Даже при таком одностороннем подходе вырисовывается очень рельефная картина, дающая основание выделить опорные места сообщения в виде существительных, глаголов и числительных для специальных текстов. Оказалось, иными словами, что при воспроизведении услышанного материала испытуемые в большинстве случаев затрудняются с ответом именно в тех местах высказывания, где выбор категорий наибольший.

Особый интерес представляло для нас соотношение информационной и грамматической нагрузок у различных элементов текста. Мы расчленили весь материал на связной текст и цифры и проанализировали его с точки зрения объема информации, воспринятой оператором.

После проведенных расчетов оказалось, что при равной энтропии источника объем воспринятой информации цифрового материала по сравнению с текстом изменяется весьма незначительно. Это наводит на мысль, о том, что на общем стандартном фоне грамматических зависимостей можно целенаправленно проводить дополнительное усиление или, наоборот, ослабление тех или иных элементов. В нашем случае — при одинаковой важности словесного и цифрового материала — такое дополнительное усиление извне получили числительные.

Таким образом, наше экспериментальное исследование показало явное наличие двух функциональных ступеней при выборном аудировании. Нижняя связана с простым физическим различением входного сообщения, более высокая и сложная связана с идентификацией слов и значений. Эту более сложную ступень можно значительно облегчить, учитывая закономерности нижнего уровня.

Подходя к идентификации слов и фраз со структурно-грамматической стороны, мы использовали разноразмерность грамматических классов и связанное с нею распределение пауз нерешительности. Совпадение данных оказалось очень показательным: наименее вероятные грамматические классы вызывают наибольшее количество пауз нерешительности.

Представляется, что структурно-грамматический подход, объединенный с информационным анализом, будет способствовать выработке инженерно-психологических рекомендаций, направленных на улучшение звеньев речевой коммуникации, а тем самым — на повышение надежности всей системы.

DISCUSSION

Wiede:

Если распознаваемость речевого сообщения связана с морфологической и синтаксической структурами данного языка, то возможности распознавания в отдельных языках с различными структурами должны быть разные. Есть у Вас относительно этого вопроса какие-либо данные?

Луцижина:

Аналогичные выводы получены немецкими психологами на материале немецкого языка. Результаты их исследований сводятся к тому, что речевое развитие — это процесс сокращения семантико-синтаксической энтропии. Такие элементы языковой структуры как подлежащее, сказуемое и дополнение по мере развития ребенка представляют для него все уменьшающуюся синтаксическую информацию, тогда как различные виды обстоятельств и определений синтаксически называются наиболее информативными.

ЛАБИАЛИЗОВАННЫЕ ФОНЕМЫ В СОВРЕМЕННОМ УЗБЕКСКОМ ЯЗЫКЕ

А. МАХМУДОВ*

В теоретической фонетике современного узбекского языка существует ряд спорных вопросов, к числу которых относится фонематичность или факультативность лабиализации гласной фонемы [ɔ], и это объясняется отсутствием экспериментально-фонетических данных.

Решение указанного вопроса потребовало объективного метода экспериментального исследования работы органов речи и состава физических свойств фонемы [ɔ]. В частности, была использована комплексная методика исследования звукового состава языка, разработанная в ЛЭФИПР I МГПИИЯ, включающая метод рентгенографический, метод определения основных физических характеристик речевого сигнала, метод спектрографический, метод аудирования.

Рентгенографическое исследование показало, что узбекская фонема [ɔ] является глубоко задним, лабиализованным гласным низкого подъема и ее огубление находится в связи с общей конфигурацией надставной трубы при ее произнесении, в частности, с особенно глубоким и задним положением языка при артикуляции этой фонемы. Это специфический гласный узбекского языка. В тюркских языках исследователи отмечают наличие этой фонемы в башкирском (Н. К. Дмитриев), татарском (В. А. Богородицкий), якутском (Е. И. Урбятова) языках.

Акустический спектр узбекского гласного [ɔ] содержит три полосы резонансных частот:

$РЧ_1$ — 300—666 гц

$РЧ_2$ — 806—1119 гц

$РЧ_3$ — 2300—3937 гц

По данным рентгенограмм, при произнесении гласного [ɔ̞] рот открыт меньше, чем при произнесении гласного [ɔ]; губы напряжены, сильно удалены от зубов и вытянуты вперед; между ними — овальное отверстие.

* ЛЭФ института языка и литературы им. А. С. Пушкина АН УзССР. Ташкент.

Акустический спектр узбекского гласного [y̤] содержит три полосы резонансных частот:

РЧ₁ — 200—666 гц

РЧ₂ — 806—1295 гц

РЧ₃ — 1902—2937 гц

При произнесении гласного [y] между губами образуется очень небольшое, почти круглое отверстие. Как показывает киносъемка, при произнесении [y] губы более собраны и более активны, чем при артикуляции [y̤].

Акустический спектр узбекского гласного [y] содержит две полосы резонансных частот:

РЧ₁ — 200—666 гц

РЧ₃ — 1687—2652 гц.

Рентгенографическое исследование позволило уточнить, что гласный [y̤] обладает более задним образованием, чем [y]. Расстояние между корнем языка и задней стенкой глотки при произнесении [y̤] составляет 26 мм, а при произнесении [y] — 29 мм.

Киносъемка губ показала, что по степени огубленности наиболее лабиализованным гласным является (y), наименее лабиализованным гласный (o), среднее положение занимает гласный (y̤). Гласные более высокого подъема аудиторы воспринимали как более лабиализованные и наоборот, гласные низкого подъема они воспринимали как менее лабиализованные.

В узбекском литературном языке гласные заднего ряда являются лабиализованными.

В зависимости от консонантного окружения возникают более передние (к |y̤| кам, к |y| лол) и более задние (к |y| лок, к |y̤| пол) варианты гласного |y̤| и |y|, не имеющие фонематического значения. В узбекском языке появление этих оттенков исторически восходит к старотюркскому вокализму, когда фонематически различались передние и задние сингармонистические параллели. На современном этапе фонологического развития это явление исчезло из узбекского литературного языка, сохранившись лишь в виде оттенков звуков речи. В андижанском, ферганском и опском диалектах узбекского языка указанные оттенки фонем |y̤| и |y| носят фонематический характер.

Наша характеристика лабиализованных гласных звуков узбекского языка в общем совпадает с описаниями крупнейших тюркологов и фонетистов — Е. П. Поливанова, А. К. Боровкова, К. К. Юдахина, Е. В. Севор-тяна, А. Н. Кононова, А. Г. Гулямова, Ф. А. Абдулаева, Ш. Ш. Шаабдурахманова, И. А. Киссена, но обоснована экспериментально.

ЛИТЕРАТУРА

1. Артемов, В. А.: *Экспериментальная фонетика*, Москва, 1956.
2. Богородицкий, В. А.: *Этюды по татарскому и тюркскому языкознанию*, Казань, 1934.
3. Дмитриев, Н. К.: *Грамматика башкирского языка*. Москва, 1948.
4. Зиндер, Л. Р.: *Общая фонетика*. Ленинград, 1960.
5. Кононов, А. Н.: *Грамматика современного узбекского литературного языка*. Москва, 1960.
6. Махмудов, А.: *Словесное ударение в узбекском языке*, Ташкент, 1960.
7. Поливанов, Е.: Звуковой состав ташкентского диалекта, „*Наука и просвещение*“, 1920, № 1.
8. Севортян, Э. В.: Материалы по сравнительной фонетике турецкого, азербайджанского и узбекского литературного языков. Сб. *Материалы по сравнительной грамматике тюркских языков*, I, Фонетика, Москва, 1955.

AKZENTOPPOSITIONEN IN DEN SERBOKROATISCHEN DIALEKTEN: ZUM 5-AKZENTSYSTEM DER SLAWONISCHEN MUNDARTEN DES ŠTOKAVISCHEN

I. MAHNKEN—J. MATEŠIĆ

In auditiver Hinsicht besteht der Unterschied zwischen den *neuštokav.* (sogen. „steigenden“) *Regressionsakzenten* des Skr. und dem älteren (im Neuštokav. nicht erhaltenen) *čakavischen Akut* in einem *gleichförmigen* Ansteigen des štokav. Akzents ' gegenüber einer *sprunghaften* Veränderung der Tonhöhe des čakav. Akuts (Ivšić:

$\overset{aaa}{\swarrow} : \underline{a}/\overset{aaa}{\text{---}}$ bzw. $\underline{a}/\overset{a}{\text{---}}$; Belić: „u početku ravan, pa zatim skoči“; Ivić: „ein jäher Sprung von einem niedrigen zu einem beträchtlich höheren Niveau“). Bestimmte slawonische Maa. des Štokav. haben im Unterschied zu den anderen štokav. Maa. diesen alten Akut in ihrem prosodischen System bewahrt, sind aber — in unterschiedlichem Ausmaß — auch von der štokav. Akzentverschiebung erfaßt worden; diese Regressionsakzente zeigen den typischen neuštokav. Tonverlauf. Diese slawon. Maa. weisen deshalb heute *drei distinktive Tonverläufe bei den Längen* und ein oder zwei Akzente bei den Kürzen auf. Herr Dr. Matešić und ich haben uns an Hand von Aufnahmen aus den slawon. Dörfern Siče und Magić Mala (mit Tonhöhenschreiber-Analysen und Sonagrammen) mit der Frage nach den oppositionsbildenden Kriterien befaßt.

Unsere Untersuchungsergebnisse bezüglich des *slawon.* Akuts decken sich mit meinen älteren Ergebnissen bezüglich des čakav. Akuts; zugleich bestätigte sich unsere Auffassung darüber, welchen *akustischen* Komponenten die Relevanz in der Konstituierung der neuštokav. Akzentopposition zukommt. Wegen der dominierenden Rolle der *Satzintonation* können weder der steigende oder fallende *Verlauf* noch die relative Lage der *nachtonigen* Silbe als invariante Kriterien für diese Tonverlaufopposition angesehen werden. In *auditiver* Hinsicht erweist sich die Opposition der neuštokav. Akzente vielmehr als eine Opposition zwischen starken und schwachen Akzenten. Die diese Opposition fundierenden *akustischen* Komponenten sind die *Abgehobenheit* bzw. der sich in den Zusammenhang *einfügende Verlauf der Frequenzmodulation*. Allen *neuštok.* Akzenten gemeinsam ist das Merkmal der *Kontinuität* der Frequenzmodulation.

Bezüglich des 5-Akzentsystems der *slawon.* Maa. (Abb. 1 u. 2) kann und muß hier nur auf den *Akut* ~ näher eingegangen werden. P. Ivić und I. Lehiste haben kürzlich in einer Anmerkung (Zbornik za fil. i lingv. 8,75) die Auffassung geäußert, daß den Akzenten ' und ~ des Slawon. *gleicherweise* ein steigender oder ebener Tonverlauf zukomme und die Opposition zwischen ihnen *nur* in der relativ zur Akzentsilbe

höheren (nach ') oder tieferen (nach ~) Lage der nachtonigen Silbe bestehe. Diese Auffassung ist nach unseren Untersuchungen unhaltbar. Die Satzintonation kann ebenso nach ~ eine höhere wie nach ' eine tiefer liegende Silbe folgen lassen, so daß die Relation zur nachtonigen Silbe auch hier *nicht* das invariante Oppositionsmerkmal sein

Die Akzente ^, " , ' , \ in Magić Mala:

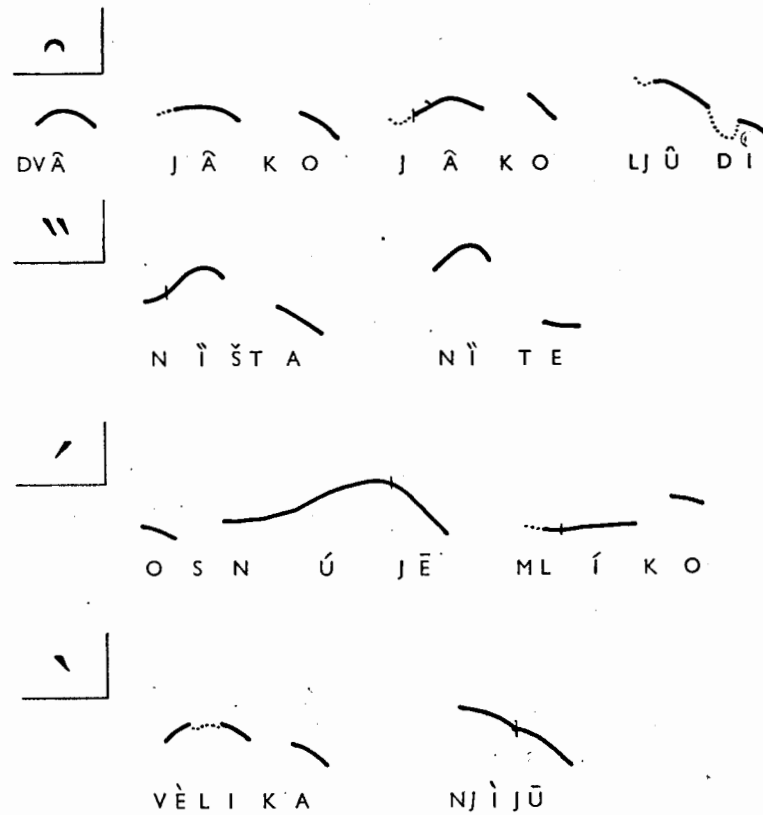


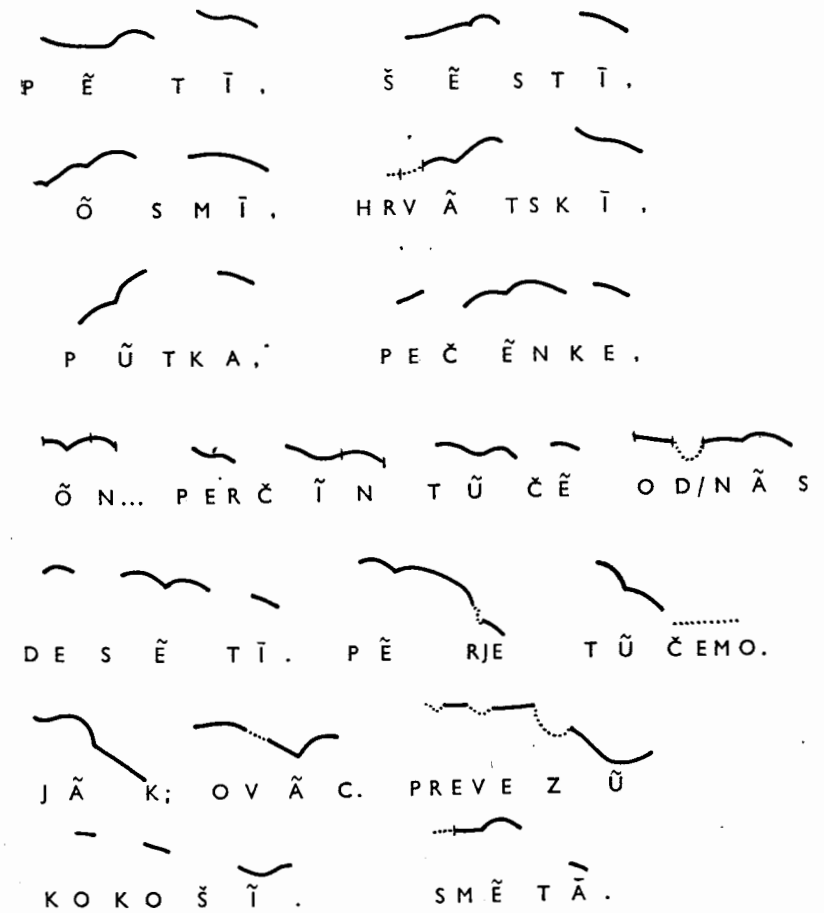
Abb. 1.

kann. Dagegen wird seit über 50 Jahren *übereinstimmend* das *auditive* Oppositionsmerkmal kontinuierlich : diskontinuierlich konstatiert. Unsere Untersuchungen galten deshalb insbesondere der dem Generalthema dieses Kongresses entsprechenden Frage: Welche *akustischen* Komponenten rufen den *auditiven* Eindruck dieser *diskontinuierlichen* Veränderung hervor?

Zunächst wurde die *Modulation der Grundfrequenz* (mit dem Tonhöhenstreiber) analysiert. Wie beim čakavischen Akut so erfolgt auch beim slawon. Akut gewöhnlich auf eine anfänglich gleichförmige (u. U. ebene) Modulation der Grundfrequenz ein *Umbruch* in eine steiler ansteigende, oft stoßtonartig auslaufende Frequenzmodulation. Beim Akut als Hauptakzent einer terminativen Kadenz liegt durchweg eine spiegelbildartige Realisierung (d. h. ein Umbruch aus einem steileren Abfall in einen

flachen, u. U. sogar steigenden Verlauf der Grundfrequenz vor. Im allg. erfolgt zugleich mit dem *Umbruch in der Grundfrequenz* eine ebenso *momentane Intensitätssteigerung* (die terminative Kadenz weist auch in dieser Hinsicht das Spiegelbild auf: die auf den intensiv artikulierten ersten Teil der betr. Silbe folgende Intensitätsreduzierung

Akut in Magić Mala: (in unterschiedlichen syntaktischen Positionen)



Čakavischer Akut: (dargestellt ist nur die Akzentsilbe bzw. deren Vokal)



Abb. 2.

betrifft insbesondere die höheren Formanten, die in dieser Position völlig aus dem Spektrum verschwinden). Spezielle Untersuchungen galten solchen Realisierungen des Akuts, bei denen die Modulation der Grundfrequenz keinen entsprechend deutlichen Umbruch aufwies. Die Sonagramme (Abb. 3—6) zeigen auch in diesen Fällen die

Realisierung des invarianten Merkmals der Diskontinuität: eine ausgeprägte gegenseitige Absetzung von zwei Segmenten, sei es durch eine momentane Veränderung der Intensitätsverhältnisse, sei es durch merkliche Veränderung im gegenseitigen Verhältnis (der Lage, dem Umfang, der Intensität) der einzelnen Formantbereiche (z. B. größere Intensität und Kompaktheit in den höheren Formantbereichen), sei es auch durch eine momentan einsetzende Verstärkung von Obertönen, die im ersten Segment kaum in Erscheinung traten. Die hier kurz skizzierten verschiedenen Arten einer momentanen Veränderung im Spektrum erscheinen u. a. in unterschiedlicher Weise miteinander kombiniert und zwar auch in Verbindung mit der eingangs behandelten Diskontinuität in der Verlaufsrichtung der Grundfrequenz. Wie meist im Bereich der prosodischen Komponenten handelt es sich somit durchweg um eine Bündelung mehrerer Merkmale. Infolgedessen kann unter bestimmten Bedingungen das übliche Hauptmerkmal durch die Realisierung anderer, unter anderen Bedingungen vielleicht redundanter Merkmale kompensiert werden. Wir kommen somit zu dem Ergebnis, daß es sich bei der Opposition der slawon. Akzente ' : ~ sowohl in auditiver als auch in akustischer Hinsicht um eine Kontinuitätsopposition (kontinuierlich : diskontinuierlich) handelt.

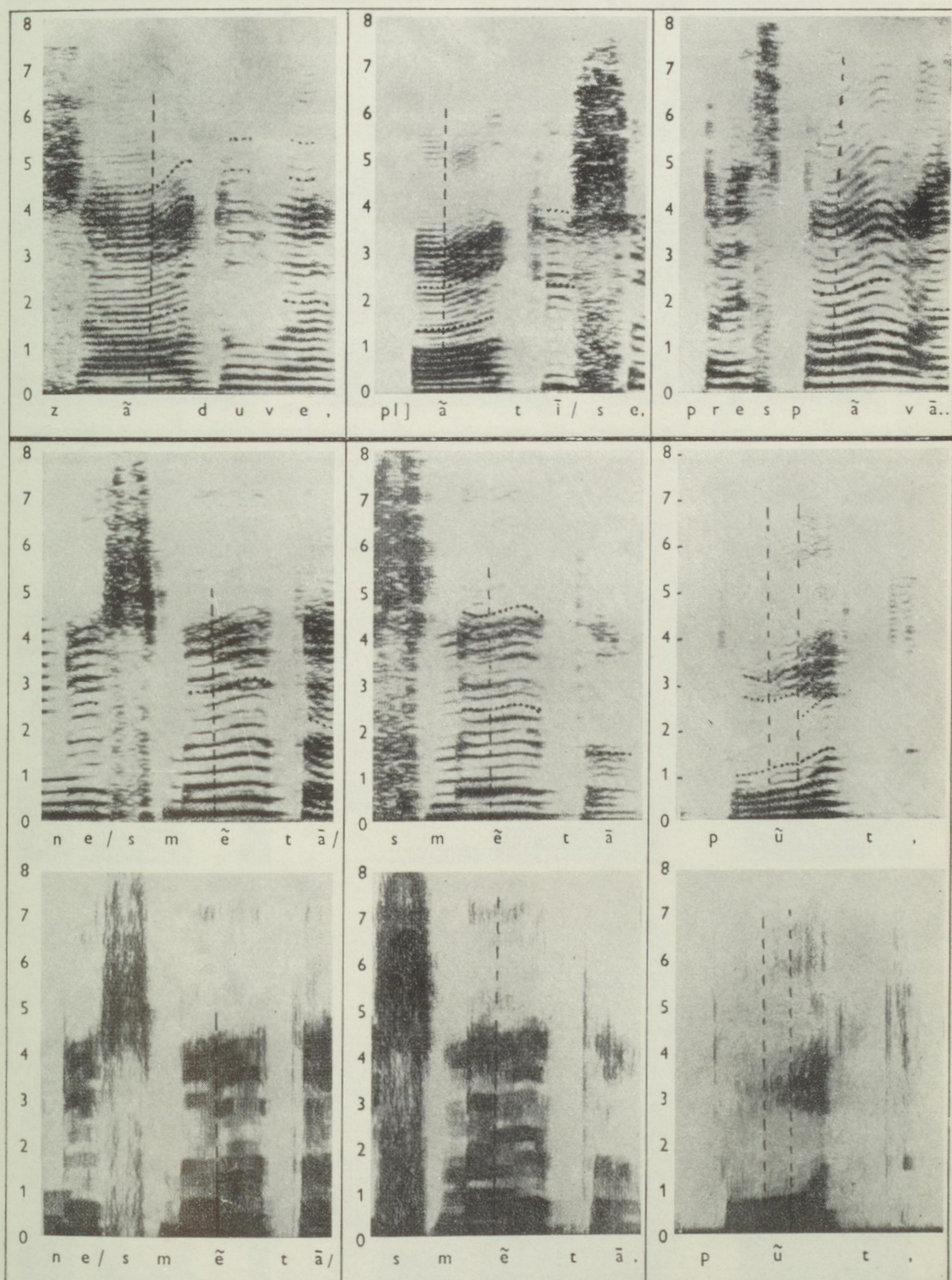


Abb. 3.

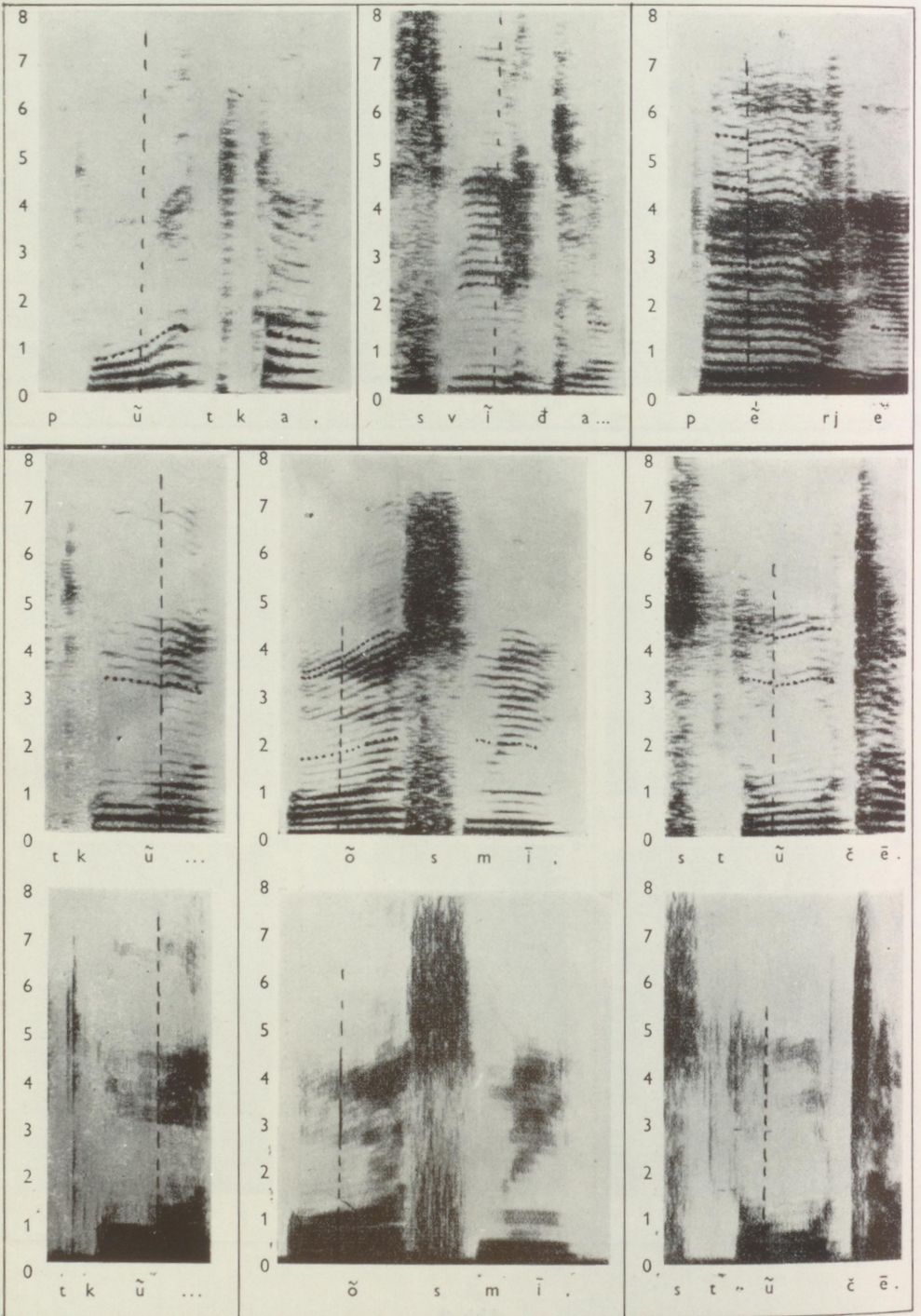


Abb. 4.

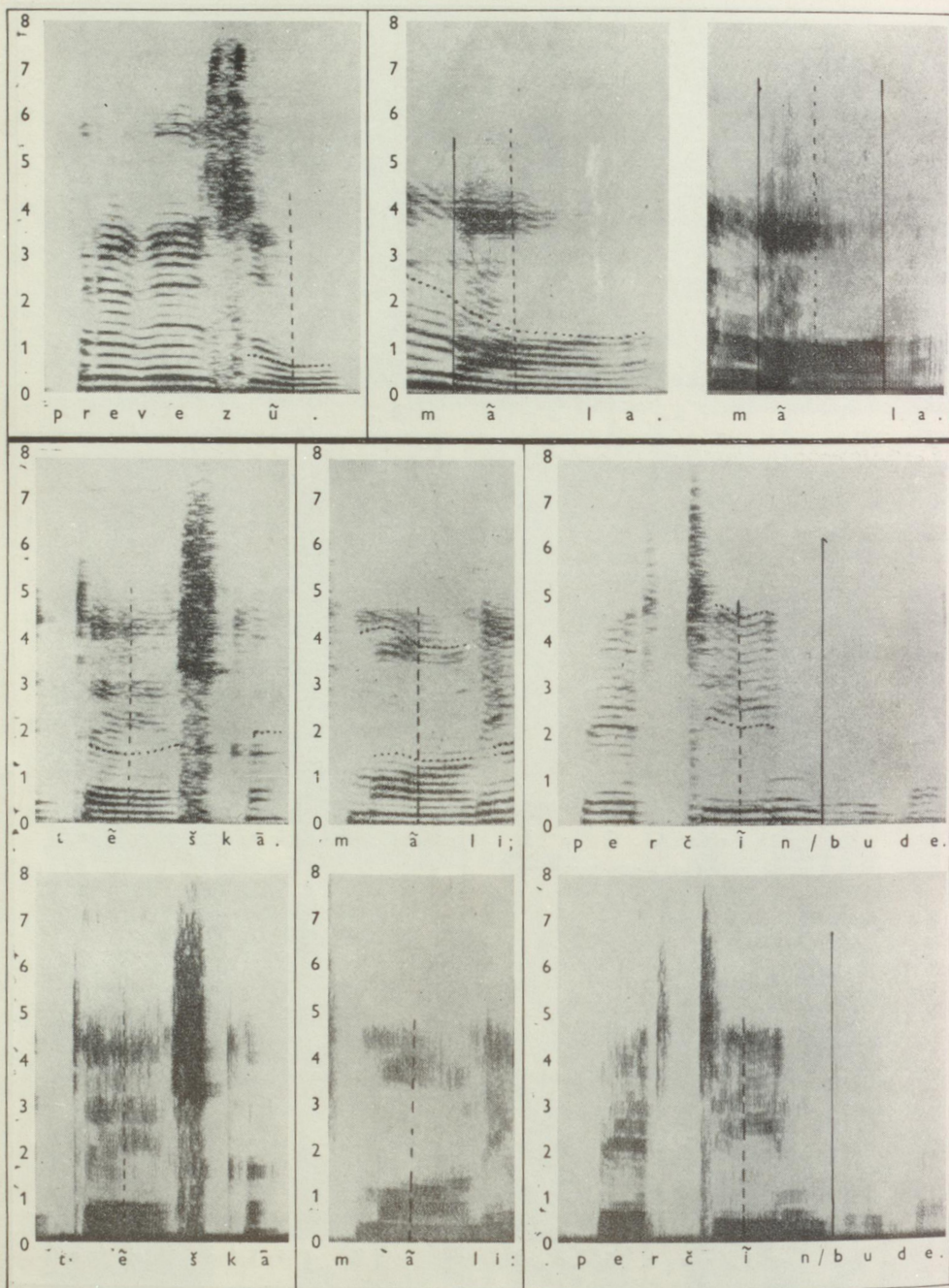


Abb. 6.

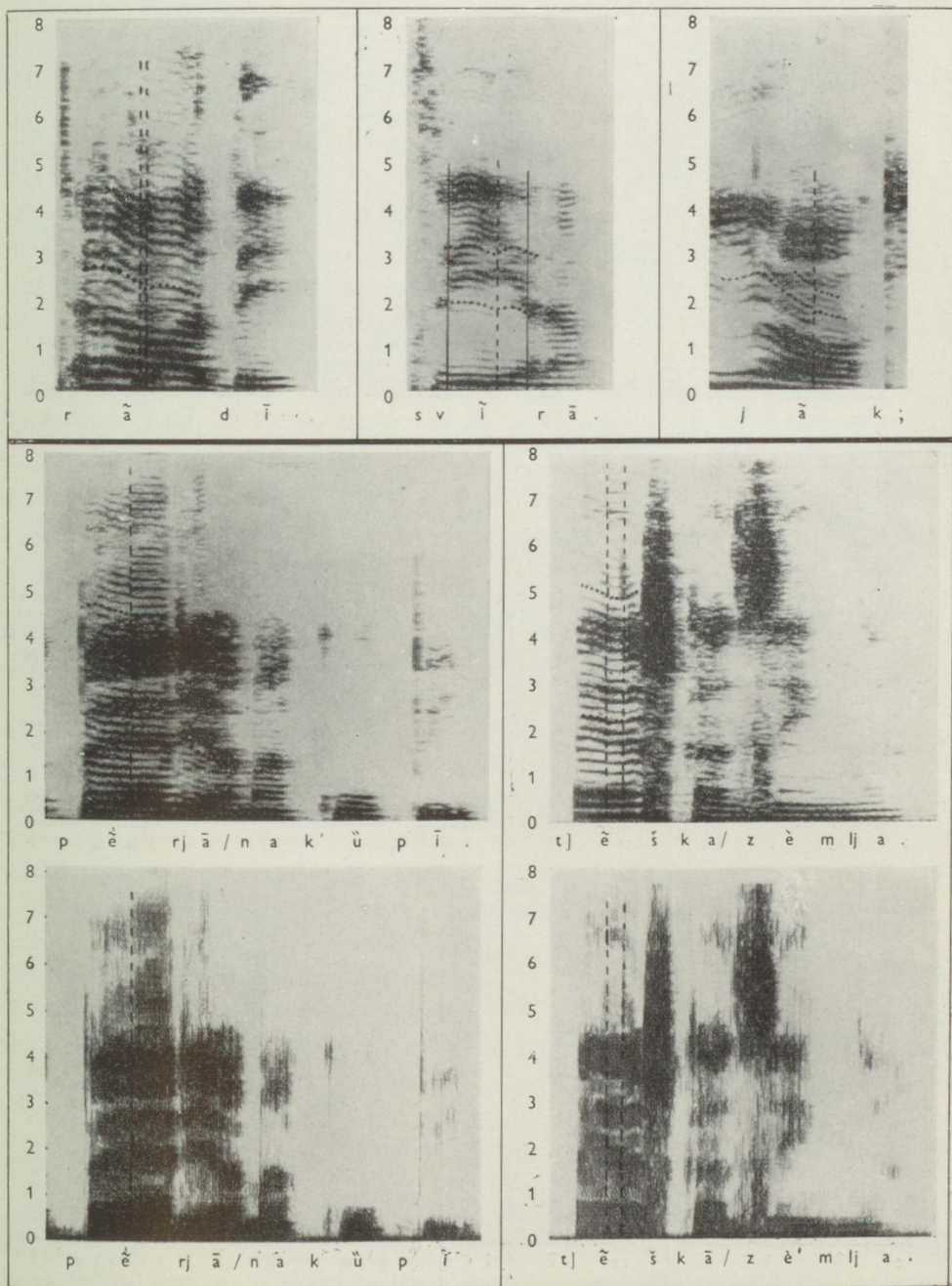


Abb. 5.

SOME PERCEPTIONAL CUES OF CZECH CONSONANTS

V. MALÁČ

My task is to relate the second part of our listening test results. The results and the conclusions from the evaluation of vowels, the first part of our test results, has already been explained by my colleague Mrs Borovičková.

The perceptual regions of relevancy for Czech consonants are not as unambiguously determined as for vowels, for which the relevancy is concentrated in the changes of frequency dimension of the speech spectra. This is caused by the fact that consonant spectra are formed by the changes of all three dimensions of the acoustic signal. Therefore the relevancy of consonants is formed also in a more complex manner. From the speech production view we may consider consonants, according to our hypothesis, as certain accurately defined distortions of the vowel sequence articulation. Even when our test was carried out mainly for the determination of relevant regions in the frequency dimension, we can determine even the relevancy of the other two dimensions.

According to our results we divided Czech consonants into three main groups. The criteria of selection were: first, those consonants which have common or similar relevant regions, second, those consonants which are most often interchanged with one another.

The first group is formed by fricatives, affricates and the consonant /ʃ/. In the second group are stops and in the third are nazals, liquids and /j/.

The characteristic feature for the first group is the noise part of the spectrum. Single consonants are differentiated by the relevancy region in the frequency dimension. The affricates are differentiated also by the duration of the sound; the pause of phonation is not very important for recognition. Voicedness is not formed by the periodical part of the spectrum alone. At high frequencies the voicedness is evidently replaced by the noise modulation in the rhythm of the pulses of the vocal cords. The relevancy of low frequencies for the recognition of voiceless fricatives is very striking. Recognition in the middle frequency bands is perhaps caused by the transient of the second formant of a following vowel.

The common trait of the second consonant group, the stops, is the great importance of pause before explosion. As the counterpart to the previous consonant group the periodical part of the spectrum has a great relevancy for the recognition of stops. The relevancy of the explosion spectrum is not equal for all members of the group. The relevancy rises in the stops sequence (p, t, t̚, k). Similarly the relevancy of the

transient of the second formant of a neighbouring vowel is not constant. According to our findings from the listening test the following vowel has greater influence on recognition of stops than the preceding vowel. This finding is in agreement with e.g. MacNeilage.

The last group, nazals, liquids and the consonant /j/ has the common feature the step change of the formant structure. The formant structure of a single member of the group is different for different sounds.

The results, on the whole, of the first and the second part of our listening test (both vowels and consonants) confirm our attitude towards the minimal unit of the acoustic speech signal. This unit is the consonant vowel combination. We do not even believe in the existence of the phoneme as a minimal unit of the acoustic speech signal.

DE FERDINAND DE SAUSSURE À ROMAN JAKOBSON— L'ARBITRAIRE DU SIGNE ET LA SUBSTANCE PHONIQUE DU LANGAGE

BERTIL MALMBERG

On sait que la célèbre thèse de l'arbitraire du signe du maître genevois a deux aspects profondément différents l'un de l'autre. D'un côté il s'agit de savoir si le signifié *boeuf* est plus motivé comme expression d'un contenu 'boeuf' que, par exemple *ochs* ou *buey*. De l'autre, il y a le problème de la délimitation d'une portion (unité discrète) du contenu (l'animal en question est un boeuf, non pas une vache ou un taureau) à l'intérieur du continuum indivisible que constitue le monde extérieur, et ensuite le problème correspondant à la délimitation analogue des unités phonologiques, par définition discrètes aussi, à l'intérieur du monde physique des sons, en principe de nombre illimité. Seul le premier de ces deux aspects retiendra notre attention ici.

Les formations imitatives et expressives — courantes dans toutes les langues avec, toutefois, une fréquence variable — sont des exemples de combinaisons motivées de signifiants et de signifiés. Personne n'appellerait le chat 'vov-vov', ni le chien 'miau-miau'. Mais même ces formations si nettement caractérisantes sont loin d'être des imitations exactes des sons symbolisés. Elles représentent en réalité une adaptation à un système phonologique de phénomènes extra-linguistiques et par là une généralisation arbitraire, socialement déterminée, de caractéristiques chez l'espèce en question qui ont été choisies à l'exclusion d'autres. Déjà Otto Jespersen avait fortement souligné (dans son compte rendu du „Cours“; réimprimé dans „Linguistica“, 1933) ce qu'il y avait d'exagéré dans la thèse du maître genevois. J'ai attiré ailleurs l'attention sur la structure phonologique pauvre de ces formations imitatives — pauvreté qui forme du reste souvent un contraste frappant avec la richesse acoustique de beaucoup de ces modèles symbolisés (cf. mes contributions aux congrès des linguistes de 1962 et des sciences phonétiques de 1964 et mon article dans *Phonetica* XI, 1964, pp. 221—227). A part ces quelques exemples de formations imitatives et expressives — courantes mais marginales dans les langues d'aujourd'hui (pour les signes appelés par Saussure „relativement motivés“, voir plus loin) — le choix de la substance sonore utilisée comme manifestation physique de la structure de l'expression (signifiant) du signe est chez Saussure et chez tous les linguistes qui se réclament de lui arbitraire (c'est-à-dire dû au hasard, ou plus exactement à une évolution historique qui, pourtant, ne se laisse pas percevoir dans l'état synchronique où fonctionne la langue en question). Ce caractère arbitraire est le même s'il s'agit du choix des phonèmes et de

leur arrangement à l'intérieur des limites admises par les règles syntagmatiques, ou du choix des faits physiques (vibrations sonores et mouvements physiologiques) utilisés pour les manifester. Je n'ai pas, en faisant cette dernière remarque, voulu prendre position ici quant au problème de savoir s'il y a un rapport quelconque entre la substance sonore et la forme que celle-ci manifeste.

La dite double articulation du langage, qui est un trait distinctif de celui-ci et qui le distingue de tous les autres systèmes de communication inventés par l'homme ou par les animaux, a été considérée comme un procédé de réaliser une expression suffisamment riche et différenciée pour correspondre aux nombreuses nécessités différenciatrices de l'homme civilisé. Les éléments phoniques passent d'un état symbolique primitif à un état uniquement différenciateur avec le passage de la communication expressive et appellative du pré-homme à la communication essentiellement énonciative des porteurs d'une civilisation plus avancée. Martinet a eu raison de souligner que c'est la deuxième articulation (l'idée des figures dans le sens de Hjelmslev) qui protégera l'expression du danger imminent du symbolisme sonore. Le langage humain correspond à des besoins communicatifs beaucoup trop compliqués pour être basé essentiellement sur un symbolisme acoustique ou — plus généralement — sur un parallélisme même approximatif entre la substance du contenu et celle de l'expression. Je reviens tout à l'heure à ce dernier point. J'ajoute d'abord que la survivance, dans les langues d'aujourd'hui, de types phonologiques à valeur expressive et symbolique dans certaines fonctions de la communication, dans certains types de vocabulaire et dans certaines formations, nous donne un exemple parmi d'autres de la conservation de systèmes primitifs à l'intérieur de structures plus complexes et plus développées. Le symbolisme phonique et le caractère arbitraire du signe, qui théoriquement sembleraient s'exclure, sont l'un et l'autre des caractéristiques du langage humain. Grâce à ses investigations poussées et à son génie synthétique et généralisateur, Roman Jakobson a pu ajouter une nuance qui manquait dans le schéma rigide du maître de Genève. L'arbitraire du signe et le symbolisme phonique représentent deux points extrêmes dans le fonctionnement du langage.

Mais revenons au problème de la substance sonore de l'expression et confrontons l'idée du symbolisme phonique avec une autre idée fondamentale saussurienne, celle de la forme. La langue est une forme, non pas une substance, selon Saussure. Et nous savons que c'est en tirant de cette thèse les conséquences les plus extrêmes que Louis Hjelmslev est arrivé à établir son système glossématique. Pour la phonologie de Prague, les phonèmes étaient des sons, à valeur distinctive, il est vrai, mais néanmoins des sons. Et c'est en tant que sons que les phonèmes possèdent certaines qualités physiques („traits pertinents“, „relevante Eigenschaften“) et grâce à ces qualités substantielles que le fameux système jakobsonien des traits distinctifs („distinctive features“) a pu être conçu. Les traits distinctifs ne figurent pas dans la théorie glossématique pour la simple raison que les unités vides („cénèmes“) qui chez Hjelmslev prennent la place des phonèmes, manquent, par définition, de caractères positifs (étant des entités „oppositives, relatives et négatives“; Saussure).

J'ai eu tout récemment, au congrès des linguistes de Bucarest, dans un contexte apparemment tout autre, l'occasion d'attirer l'attention sur la place des formations imitatives et expressives (onomatopées, interjections, langage enfantin, etc.) en marge du système conventionnel du langage et de souligner que, plus ces formations sont indépendantes des normes phonologiques, moins elles participent aux mutations. Elles ont l'air d'être indépendantes des dites lois phonétiques. Des types comme *ɣst*, *chut*, n'ont pas d'histoire. Ils sont dus directement à leur valeur imitative ou expressive. L'évolution du langage est une modification des normes sociales qui déterminent la structure linguistique. Une formation qui est indépendante (en principe) de telle ou telle tradition phonologique, reste donc invariable. Ou, pour le dire autrement, elle est recrée toutes les fois qu'elle est utilisée. Ce qui a été dit vaut en principe pour un cas extrême, la formation extra-phonologique pratiquement inexistante. Une telle formation, ne participant pas à la deuxième articulation, ne serait pas un signe et ne se laisserait par conséquent pas non plus scinder en contenu et en expression.

Il est vrai, d'autre part, que les symboles à une seule articulation ou dénués de structure proprement dite sont souvent, eux aussi, entièrement ou essentiellement arbitraires. Ainsi par exemple les messages des animaux. Sebeok rappelle (*Revue internationale des sciences sociales* XIX, 1967, no. 2; p. 7 du tirage à part) que les mouvements de la queue indiquent l'amitié chez le chien, l'hostilité chez le chat et la présence de mouches chez le cheval. Le développement et la complexité croissante de la deuxième articulation ne sont donc pas nécessairement parallèles à une évolution vers l'arbitraire, comme on pourrait le penser.

Je reviens à cette occasion aussi à un autre aspect de la définition saussurienne du signe, son caractère d'être articulé, par conséquent composé — en tout cas son expression — d'éléments plus petits et plus généraux: les figures de Hjelmslev, les phonèmes des phonologues, à leur tour composés des traits distinctifs de Jakobson, etc. Le principe de la deuxième articulation implique l'imprévisibilité des éléments dans la chaîne — principe qui, pourtant, n'est jamais appliqué cent pour cent mais, au contraire, fortement restreint par les règles syntagmatiques de chaque langue (règles dites de distribution: groupement des phonèmes, structure des syllabes, faits de neutralisation, d'assimilation, d'harmonie vocalique, etc.).

Si nous revenons pour un instant aux formations imitatives, enfantines, expressives etc. auxquelles nous avons fait allusion tout à l'heure, nous savons que leur structure représente aussi d'une façon générale un type plus simple, admettant moins de variations: harmonie vocalique et consonantique, combinaisons consonantiques simples ou absence de groupes, syllabes ouvertes etc., là où le système du vocabulaire normal représente des structures plus complexes. L'une des deux extrémités du langage est donc représentée par le symbole non-articulé, indivisible, où il y a identité entre contenu et expression et un rapport naturel entre les deux; l'autre par le signe saussurien pleinement développé avec sa séquence d'éléments d'expression grosso modo imprévisible l'un à partir de l'autre et de la substance du contenu. Nous voyons donc encore une fois comment différents modes de description du langage et de ses éléments

— comme arbitraires ou comme motivés, comme prévisibles ou comme imprévisibles, comme simples ou comme complexes — se justifient au même titre, vu les différentes couches qui se retrouvent toujours et partout dans toute communication humaine.

Hjemslev a développé jusqu'à l'absurdité même la thèse de la forme — pas encore entièrement utilisée chez Saussure (cf. mon article „Ferdinand de Saussure et la phonétique moderne“, *Cahiers Ferdinand de Saussure* 12, 1954). Il est donc allé jusqu'à vouloir exclure la substance du domaine de la linguistique. En ce faisant, il a coupé d'une façon absolue tout lien substantiel possible entre expression et contenu. Il n'y a pas de place pour un symbolisme phonique dans un tel système. Le seul genre de motivation du signe qui reste possible serait celle qui est basée sur l'appartenance à un même paradigme (*apprendre, apprenti, apprentissage*; cf. ci-dessus) dont on connaît le rôle dans le fonctionnement communicatif du langage („rapports associatifs“ de Saussure) et qui avait figuré dans les tentatives étymologiques des précurseurs de la linguistique proprement dite (voir par exemple Leroy, „Les curiosités linguistiques de Leibniz“, *Revue internationale de philosophie* 76—77, 1966, p. 197).

Jakobson a pris la route opposée. Il a vu que le symbolisme phonique joue un rôle plus grand — même à des niveaux supérieurs du langage — qu'on n'avait voulu l'admettre. C'est par ses analyses de certaines formes de langue réduites (enfantines, aphasiques d'un côté, expressions imitatives et poétiques de l'autre) qu'il est arrivé à décrire avec une méthode exacte qui manquait aux précurseurs les traits généraux des structures qui déterminent les formes utilisées sur le niveau expressif du langage. Jakobson nous a montré la route. On n'a qu'à suivre sur ses pas. Il ne fait pas de doute qu'il y a des découvertes intéressantes à faire dans ce domaine. Il s'agit pourtant d'abord d'objectiver les méthodes. La psychologie de la perception possède maintenant les procédés et la technique à utiliser. Les phonéticiens n'auront qu'à les appliquer.

Cette nouvelle branche de la phonétique — qui suppose, inutile de le souligner, une prise en considération constante et, dans ce cas, particulièrement difficile du contenu (comme forme et comme substance) — n'exclut pas, et se laisse parfaitement bien réconcilier avec une conservation et un développement d'autres lignes de recherches plus formelles, d'orientation structurale et même directement glossématique. L'arbitraire du signe et la motivation du signe ne s'excluent donc pas. C'est entre ces deux pôles que se réalise le langage du genre humain, et c'est entre ces deux pôles aussi que, par conséquent, doivent se développer les activités scientifiques du linguiste et du phonéticien.

DISCUSSION

Akhmanova:

The problem of the arbitrary character of the linguistic sign has been discussed time and again, but very much still remains to be done. One would, therefore, be pleased to find that new appro

aches to the problem are now being discovered. I would even go so far as to suggest that a "revision" of the too rigid original formula may at present be called for.

Holt:

Il y a dans la glossématique une place réservée aux phénomènes comme les interjections (dans *pst, hm*), si l'on prend en considération l'existence du syncrétisme. Pour les interjections véritables, il y a syncrétisme sur deux plans linguistiques, celui de l'expression et celui du contenu.

Martinet:

Il n'est pas possible de définir ce qu'est une langue sinon en fonction de l'arbitraire tel qu'il se dégage de l'enseignement de Saussure. Ce qui est proprement humain dans le langage est précisément ce qui est arbitraire. Une fois ceci acquis, on abordera évidemment avec fruit les éléments expressifs marginaux.

Mattoso—Camara:

Je demande pardon de dire quelques mots sur ce sujet. J'ai médité beaucoup sur le principe de Saussure et l'apparente contradiction qui découle de ce qu'on rencontre des onomatopées et des sons expressifs. Je crois qu'on pourrait relier ces deux faits avec les fonctions du langage établies par Bühler: *Kundgabe, Appel, Darstellung* ou plutôt, *Bericht* „information“, puisque toutes ces fonctions sont représentatives. L'arbitraire s'affirme dans l'information, quand le langage passe à ignorer les valeurs expressives qui sont nécessairement incluses dans le mot. Aussitôt qu'on passe aux autres fonction (p. ex. le langage poétique) ces valeurs expressives montent à la surface et on a des sons imitatifs ou expressifs.

Vachek:

Also the primitive elements are, in a way, arbitrary, and thus integrated in the language system. This may be proved by instances of doublets like Czech *čichat — čuchat*, the latter word being stamped as peripheral by the otherwise unusual phonemic combination *č + back vowel*.

Malmberg:

Je suis entièrement d'accord avec M. Vachek qu'il y a dans les structures linguistiques des degrés de toute sorte, et je souligne encore que toutes les formations auxquelles j'ai fait allusion se retrouvent quelque part entre les deux extrêmes: le continuum indivisible (motivé) et le complexe d'éléments discrets (arbitraire). Je suis gré à M. Holt d'avoir attiré l'attention sur le syncrétisme qu'il y a entre contenu et expression dans celui-là, et à Madame Akhmanova d'avoir appuyé en principe mon point de vue. Quant à l'identification que fait M. Martinet entre les formations imitatives et les gestes, je suppose qu'il pense aux interjections (*pst*, etc.) — pour lesquelles je suis en principe d'accord. Les gestes contenant aussi un certain élément social et arbitraire — et non pas à des mots comme *chuchoter* ou *siffler*. Cela a été un des buts de mon intervention ici d'attirer l'attention aux nombreux faits de vocabulaire phonologiquement structurés mais où les traces d'une origine onomatopéique se sont conservées et où nous avons toujours à faire à des signes partiellement motivés. L'emploi qu'en fait l'homme parlant est une preuve suffisante de leur caractère imitatif. C'est une tâche pour la phonétique de chercher ce qu'il y a de général dans leurs fonctions expressives, poétiques, etc.

DEUX LOIS DE L'ÉVOLUTION PHONÉTIQUE

WITOLD MAŃCZAK

On peut formuler, entre autres, deux lois de l'évolution phonétique qui s'appliquent à toutes les langues.

La première, qui concerne le développement phonétique régulier, est la suivante: *l'évolution des phonèmes plus rarement employés est, en général, plus rapide que celle des phonèmes plus fréquemment usités.* Par ex., les consonnes sonores sont plus rarement employées que les sourdes, ce qui fait que les sonores changent plus vite. A la suite de la 1^{re} palatalisation, **k* est devenu, dans toutes les langues slaves, une affriquée (*č* ou *c*), tandis que **g* a abouti, également partout, à une affriquée (*ž* ou *z*), bien entendu par l'étape intermédiaire d'une affriquée. A la suite des 2^e et 3^e palatalisations, **k* est devenu partout une affriquée (*c*), alors que **g* est devenu soit une affriquée (*g*) soit une fricative (*z*). Il en est de même pour le développement de **tj*, **dj*, qui parfois présente un parallélisme, mais, s'il y a une différence entre l'évolution de **tj* et celle de **dj*, le groupe à occlusive sonore évolue plus vite.

Cette loi n'est qu'un cas particulier d'une loi plus générale, d'après laquelle toute évolution linguistique (phonétique, analogique, sémantique) dépend de la fréquence.¹

La seconde loi, qui concerne l'évolution phonétique irrégulière, est la suivante: *les éléments linguistiques (morphèmes, mots, groupes de mots) dont la fréquence augmente diminuent, en général, de volume.* Cette loi diachronique est une conséquence logique de la loi synchronique de Zipf. Il est vrai que par la loi de Zipf on entend normalement une formule relative au rapport entre la longueur des mots et leur fréquence, mais il résulte de la lecture des ouvrages de Zipf qu'il s'est rendu compte du rapport entre la fréquence et le volume de n'importe quel élément linguistique, de sorte qu'il est permis d'entendre par la loi de Zipf une loi aussi générale que la suivante: les éléments linguistiques plus petits sont, en général, plus souvent employés que les éléments plus grands. S'il s'agit de la phonétique, on peut citer par exemple que les voyelles nasales, qui sont des combinaisons *voyelle orale + nasalité*, sont plus rarement employées que les voyelles orales et qu'il en est de même pour le rapport entre les consonnes sonores (= *sourdes + sonorité*) et les sourdes. Comme la différence entre synchronie et dia-

¹ Pour plus d'exemples v. notre communication *Słowiański rozwój fonetyczny w świetle pewnego ogólnego prawa*, Sprawozd. z posiedz. komisji oddz. PAN w Krakowie, janvier—juin 1966.

chronie est quantitative, le développement linguistique n'étant autre chose qu'une somme d'états de langue, il est évident que l'augmentation de la fréquence d'un élément linguistique doit avoir pour conséquence une réduction de son volume. On connaît des abrégements graphiques (*page* > *p.*), morphologiques (*télévision* > *télé*) et syntaxiques (*chemin de fer métropolitain* > *métropolitain*), sans tenir compte de ce qu'il y a aussi des abrégements phonétiques dus à la fréquence. Faute de connaître cette loi, on invente des explications erronées (mots atones, „Allegroformen“, „Funktionslosigkeit“, etc.) ou bien on laisse des irrégularités sans explication (*sapiō* > *sais*, *françois* > *français*).

Il y a quatre moyens de reconnaître qu'il s'agit d'un développement phonétique irrégulier dû à la fréquence:

a) S'il existe, pour la langue et la période en question, un dictionnaire de fréquence, on peut, bien entendu, en profiter, parce que la grande majorité des mots subissant un développement phonétique irrégulier dû à la fréquence se trouvent parmi les mille mots le plus souvent employés dans la langue en question.²

b) Si le morphème, mot ou groupe de mots apparaît dans une langue donnée sous une forme double, régulière et irrégulière, le développement phonétique irrégulier dû à la fréquence se caractérise par le fait que la forme irrégulière est en général plus souvent employée que la régulière, cf. *me* — *moi*, *ne* — *non*, *la* — *elle*, *monsieur* — *monseigneur*, *français* — *François*, etc.

c) A côté du développement phonétique irrégulier dû à la fréquence, il existe une autre évolution irrégulière, consistant en assimilations, dissimilations, métathèses, haplogogies, formes hypercorrectes ou expressives. Or, toute cette évolution irrégulière se caractérise par le fait qu'elle a lieu, dans différentes langues, dans les mots les plus divers. *Chercher* < *cercher* présente une assimilation, *faible* < *flébilem* une dissimilation, *troubler* < *turbulāre* une métathèse, mais il serait difficile de trouver, dans une langue indo-européenne, un mot signifiant „chercher“ qui présenterait une assimilation, un mot signifiant „troubler“ avec une métathèse. Alors qu'il n'y a aucun parallélisme entre les irrégularités dites assimilations, dissimilations, etc., au contraire le développement phonétique irrégulier dû à la fréquence s'accomplit dans différentes langues d'une manière plus ou moins parallèle (cf. l'évolution des numéraux ou bien des mots signifiant „monsieur“), ce qui s'explique par le fait que les mots les plus fréquents se rapportent tous plus ou moins aux mêmes notions.

d) En ce qui concerne les changements phonétiques irréguliers dus à la fréquence qui se produisent à l'intérieur du paradigme flexionnel d'un mot, les abrégements irréguliers ont lieu plus souvent dans les formes plus fréquentes que dans les formes plus rares. Considérons le développement du prés. de l'ind. *habeō*, *habēs*, etc. et du prés. du subj. *habeam*, *habeās*, etc. en italien. Parmi les formes *ho*, *hai*, *ha*, *abbiamo*:

² V. notre communication *Développement phonétique irrégulier et fréquence d'emploi en français*, Actes du X^e Congrès Internat. de Ling. et Phil. rom., Paris, 1965, p. 911—914.

avete, *hanno*, sont irréguliers *ho*, *hai*, *ha*, *hanno*, ce qui s'explique par le fait que le singulier est plus souvent employé que le pluriel et que la 3^e personne est plus souvent employée que les autres. Les formes du prés. du subj. *abbia*, *abbia*, *abbia*, *abbiamo*, *abbiate*, *abbiano* sont toutes régulières, ce qui s'explique par le fait que le subjonctif est plus rarement usité que l'indicatif.³

DISCUSSION

E. Novák:

Il me semble nécessaire de compléter la première loi de M. Mańczak („L'évolution des phonèmes les plus rarement employés est, en général, plus rapide que celle des phonèmes plus fréquemment usités.“) par la thèse suivante: Partout là où une subdominante existe, dans le système phonologique, à côté d'une dominante (termes pris de la musicologie), cette dominante tend à supprimer la subdominante. Par ex. dominante — subdominante *a* — *ä* en slovaque, *a* — *ü* en allemand, *a* — *â* [à] en français, *á* — *ε* en hongrois, etc. (Pour les détails voir notre étude intitulée *Dominante — subdominante en phonologie* qui paraîtra dans le III^e tome des *Travaux linguistiques de Prague* en 1968).

Vachek:

The quantitative factor is certainly important in language development but it is not the only one; it cooperates with the qualitative factors. This is shown, e.g., in the case of *Modε* [ž] which is not changed, although its frequency is very small, because it is propped up and backed by the existence of the [š]-phoneme in *Modε*. One has to consider the existence of the functional factor in language development to be at least of equal importance as the quantitative factor.

Mańczak:

L'explication de Horn est erronée. Un dictionnaire comprend souvent 40 000 mots: la plupart d'entre eux contiennent des phonèmes superflus (par ex. *faculté* sans *l* serait intelligible), un petit nombre d'entre eux subissent des réductions. J'ai établi que la plupart des mots réduits se recrutent parmi les mille mots les plus fréquents, qui constituent 2% de mots enregistrés dans un dictionnaire. Cela prouve que ces réductions sont dues à la fréquence et non à la „Funktionslosigkeit“.

³ Pour plus de détails, v. nos manuels *Phonétique et Morphologie historiques du français*. Łódź, 1962, ainsi que *Polska fonetyka i morfologia historyczna*, Łódź, 1965.

CONSIDÉRATIONS STATISTIQUES SUR LES FINALES VOCALIQUES EN ROUMAIN

CONSTANT MANECA

1. La fin du mot¹ en roumain présente un intérêt particulier parce qu'il porte souvent la marque d'importantes catégories comme le nombre, la personne, le temps, etc. Dans notre communication nous allons voir la situation actuelle des finales vocaliques de la langue littéraire et du dialecte qui est à sa base (celui de Valachie). Dans ce but nous avons fait une investigation statistique sur des enregistrements des textes littéraires (de la comédie de T. Muşatescu, *Titanic Vals*² et du volume *Întuneric si lumină* de I. Al. Brătescu-Voineşti³ et sur des textes dialectaux de la Valachie.⁴

2. La situation des finales vocaliques selon les données de notre statistique⁵ est la suivante:

Son	Prononciation littéraire				Prononciation dialectale	
	T. M.	B.—V.	Total	%	Total	%
—a	1 457	1386	2 843	14.215	3 309	16.545
—e	2 099	2178	4 277	21.385	1 958	9.79
—i	513	674	1 187	5.935	1 437	7.185
—o	384	248	632	3.16	731	3.655
—u ⁶	760	878	1 638	8.19	1 311	6.555
—ă	1 269	978	2 247	11.235	3 769	18.845
—î	5	19	24	0.12	491	2.455
—î ⁷	1 431	957	2 388	11.94	1 630	8.15
—ă ⁷	251	60	311	1.555	919	4.595
Total	8 169	7378	15 547	77.725	15 555	77.775
consonnes	1 831	2622	4 453	22.265	4 445	22.225

¹ Pour notre recherche nous avons délimité le mot selon le *Dicţionarul limbii române moderne* (Bucureşti, 1958) de l'Académie de la R. S. R., comme l'ont fait d'ailleurs MM. A. Juilland, P. M. H. Edwards et Mme I. Juilland dans leur *Frequency Dictionary of Rumanian Words* (London, The Hague, Paris, 1965).

² D'après un enregistrement d'une émission de théâtre au microphone de la Radiotélévision Roumaine.

³ Nous avons choisi cet auteur parce qu'il était un des sujets de l'enquête de l'ALR. L'enregistrement a été fait d'après la lecture d'un chercheur de l'Institut de linguistique, originaire de la région d'Arges.

⁴ Emil Petrovici, *Texte dialectale*, Sibiu—Leipzig, 1943.

⁵ Le corpus de notre statistique est composé de: 10 000 occurrences de chaque auteur et 20 000 occurrences des textes dialectaux.

3.1. La proportion des finales vocaliques dans les deux aspects de la langue roumaine est relativement la même. Selon l'ordre de la fréquence les finales vocaliques se distribuent de la manière suivante:

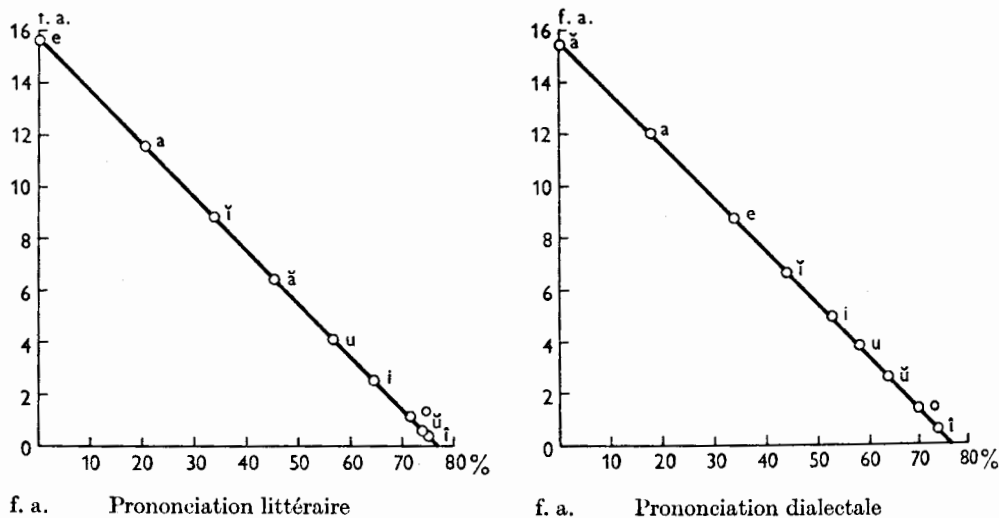


Fig. 1.

3.2.1. Première observation: le flux de la communication du roumain présente d'habitude des finales vocaliques (13.095 dans la prononciation littéraire et 13.006 dans la prononciation dialectale valaque, c'est à dire 65,475 % et 65,03 % du total des finales des mots) ou semi-vocaliques (2452 dans la prononciation littéraire et 2549 dans la prononciation dialectale de la Valachie, c'est à dire 12,26 % et 12,726 % du total des finales) ce qui contribue à une meilleure audibilité de la fin du mot.⁸

3.2.2. Seconde observation: dans les deux variantes de la langue la proportion entre les finales vocaliques, semi-vocaliques et consonantiques varie très peu, donc la variation est pratiquement négligeable.

3.3.1. En comparant la situation des finales vocaliques de la prononciation littéraire et celle dialectale de la Valachie nous observons que seulement -a et -i gardent leurs postes dans l'échelle des fréquences, présentant pourtant d'importantes différences aux valeurs absolues et relatives. La distribution quantitative des voyelles finales du roumain littéraire est bien différente de celle du dialecte valaque. C'est

surtout à cause du remplacement réciproque (déterminé dans le dialecte par la qualité dure ou palatale de la consonne précédente) qui peut intervenir entre -ă et -e, qui apparaissent à la fin des mots fréquents comme les prépositions *de*, *pe*, le pronom réfléchi *se*, etc.⁹ On peut observer ici l'influence de la fréquence de certains mots sur la structure phonétique du langage.

3.3.2. Notre recherche vient d'apporter aussi une certaine précision quantitative sur le remplacement réciproque de l'ă et de l'e (en position finale bien entendu) qui est nettement en faveur de l'ă.¹⁰

3.4. Enfin, cette statistique nous montre une fois de plus la différence entre l'aspect écrit des finales vocaliques du roumain et la manière dont elles sont prononcées.

DISCUSSION

Mme Gsell:

Beaucoup de chercheurs se penchent en ce moment sur le problème de l'étude statistique des phonèmes pour comparer des systèmes. M. Maneca a ainsi opposé le roumain littéraire au parler dialectal de la Valachie. M. Copszinsky a étudié au 10^e Congrès de Linguistique de Bucarest les divergences entre les phonèmes polonais en toute position dans les vers et dans la prose. J'ai moi-même présenté une communication portant sur la prose et la poésie française. On fait des travaux sur l'anglais (prose et poésie) à l'Institut des langues étrangères de Moscou, et je serais seulement désireuse de savoir si des études de ce genre ont été entreprises en Roumanie sur l'étude statistique des phonèmes de la prose et de la poésie, car il serait intéressant de voir si l'idée exprimée par M. Jacobson pourrait être vérifiée, à savoir que le système phonologique de la langue est renforcé en poésie.

⁸ Il y a ici même les finales orthographiées -ul pour lesquelles nous n'avons rencontré que la prononciation -u.

⁹ Nous avons considéré ces sons aussi des voyelles, quoiqu'il y en a des discussions (cf. Al. Rosetti, *Introducere în fonetică*, București, 1963, p. 58 sqq. et E. Vasiliu, *Fonologia limbii române*, București, 1965, p. 101 sqq.).

¹⁰ Cf. Jespersen, *Lehrbuch der Phonetik*, Leipzig, 1904, p. 186 sqq.

⁹ Cf. FDRW, s. v. v.

¹⁰ Pour la position finale l'affirmation que *a este mai des înlocuit prin e datorită tocmai tendinței de palatalizare (des consonnes précédentes)* (cf. I. Coteanu, *Elemente de dialectologie a limbii române?* București, 1961, pg. 79) ne se vérifie pas.

OPTISCHE UNTERSTÜTZUNG AKUSTISCHER SPRACHSIGNALE

PETER MARTENS*

Zum Erlernen einer richtigen Aussprache und zum richtigen Erfassen von Lauten ist es erforderlich, das Gehörte in das phonologische System der betreffenden Sprache einzuordnen. Ein wichtiges Hilfsmittel für diesen Lehrvorgang ist die optische Unterstützung akustischer Sprachsignale. Damit ist folgendes gemeint: sowohl bei Normalhörenden als auch bei Gehörgeschädigten reicht der auditive Weg allein zum Erlernen oft nicht aus. Deshalb ist es vorteilhaft, sich nicht einfach auf die Nachahmung des Gehörten zu beschränken — sofern es die Aussprache betrifft — oder nur dem Ohr die fremdsprachlichen Gebilde mit neuen Lauten zu bieten, sondern die wichtigeren Einzelheiten der Artikulation durch Abbildungen zu verdeutlichen. Dabei ist es allerdings wesentlich, daß schon beim Hörenlernen die Differenzierung auf das phonologische System der Zielsprache ausgerichtet ist.

Die Bewegung der Lippen kann der Lehrende im allgemeinen dem Studenten ohne Hilfsmittel demonstrieren. Aber was hinter den Lippen und den Zähnen vorgeht, ist nur zu erfassen, wenn zwischen Lippen und Zähnen ein genügend großer Einblick frei ist. Diese Voraussetzung aber trifft eben sehr oft nicht zu. In solchen Fällen kann ein Bild von Zungenstellung und Zungenlage über den visuellen Eindruck eine Hilfe für das differenzierende Erkennen und für das Artikulieren sein. Dem Lernenden muß es erleichtert werden, die neuen Laute in ein Verhältnis zu setzen zu anderen Lauten. Diese anderen Laute können nur bekannte Laute der Muttersprache sein oder aber Laute der Zielsprache.

Ein paar Beispiele mögen diesen Aspekt verdeutlichen:

Abb. 1¹: Bei den Vorderzungenvokalen [ε] bis [ɪ] erkennt man ohne Schwierigkeiten, daß die Lippen gespreizt sind und daß der Abstand zwischen den Zahnreihen vom [ε] zum [ɪ] immer geringer wird.

Die Seitenrisse zeigen nun darüber hinaus, daß gleichzeitig der vordere Zungenrücken sich von [ε] zum [ɪ] immer weiter nach vorn oben hält, während der vordere Zungenrand hinter den unteren Schneidezähnen bleibt; Abb. 2. Bei den *Hinter-*

* Universität Hamburg.

¹ Diese Abbildungen sind der Diapositivreihe „Deutsche Laute im Bild“ entnommen (mit freundlicher Genehmigung des Verlags Hanex, Hamburg 22) dem Teil 4a der Phonetik der deutschen Sprache von Carl und Peter Martens (Verlag Max Hueber, München).

zungenvokalen sieht man ohne weiteres, daß die Lippen gerundet und vorgestülpt sind. Daß sich dabei aber gleichzeitig vom [ɔ] zum [u] der hintere Zungenrücken immer weiter nach hinten hebt, kann der Lernende nicht von außen erkennen. Dies läßt sich nur mit Hilfsmitteln wie dem Röntgenbild oder einem solchen Seitenriß demonstrieren. Abb. 3: Das äußere Erscheinungsbild der Lippen kann aber auch täuschen: beim [u] und [y] ist es etwa gleich. Will man also den akustischen Eindruck optisch unterstützen, so helfen dabei Seitenrisse: sie zeigen die entscheidenden Unterschiede der Zungenstellung. Wenn wir den Vergleich auf [u], [y], [i] ausdehnen (wie auf der Abbildung) dann haben wir bereits etwas weiteres getan: wir haben den Laut eingeordnet in das phonologische System der Zielsprache, d. h. hier also in das phonologische System des Deutschen. Das erleichtert dem Lernenden das Zurechtfinden sowohl beim Hören als auch beim Sprechen. Dies ist um so wichtiger, je ähnlicher die erlernenden Laute etwaigen Lauten der Muttersprache (oder der defektiven Sprechweise) sind. Das gleiche gilt aber auch, wenn die zu erlernenden Laute einander ähnlich sind, vom bisher benutzten phonologischen System aus gesehen (also z. B. dem der Muttersprache). Es kommt also auf die Verdeutlichung von distinktiven Faktoren an. Abb 4: Bei [e] und [ɛ] kann man auf Grund der verhältnismäßig weit geöffneten Lippen auch von außen Unterschiede der Zungenhebung erkennen. Der wirklich entscheidende Unterschied zwischen dem gespannten *geschlossenen* Vokal [e] und dem ungespannten *offenen* Vokal [ɛ] wird aber durch Seitenriß und Palatogramm erst offensichtlich. Damit ist dann auch *ein* Prinzip des Systems gezeigt: die Opposition zwischen „geschlossen“ und „offen“. Es kommt immer wieder darauf an, den Lernenden nicht einen isolierten Laut üben zu lassen, sondern ihm bewußt zu machen, wie dieser Laut in das phonologische System eingeordnet ist.

Abb. 5: Die Differenzierung von Lauten gleicher Bildungsweise wird erleichtert, wenn die Bildungsstelle klar bewußt ist, vor allem wenn die Bildungsstelle mit ihrem Wert innerhalb des Systems der betreffenden Lautgruppe bewußt ist. Nehmen wir zum Beispiel die drei Bildungsstellen für [p] [t] [k]. Das Bild erleichtert — wie schon gesagt — dem Normalhörenden wie dem Gehörgeschädigten auch durch den Anreiz zum Mitartikulieren das Differenzieren des Gehörten. Abb. 6: Unterschiede zwischen gespannten stimmlosen Konsonanten und ungespannten stimmhaften Konsonanten werden von Deutschlernenden und auch von Gehörgeschädigten oft durch übermäßige Stimmhaftigkeit der ungespannt stimmhaften Konsonanten gekennzeichnet.

Beides (Stimmhaftigkeit wie Behauchung) läßt sich mit Fingern und Spiegel ertasten und erkennen. Die wesentlichen Unterschiede in der Muskelspannung jedoch sind normalerweise nicht so leicht zu erfassen (Demonstrationen des intraoralen Luftdrucks mit Hilfe eines Wassermanometers sind für praktische Zwecke im allgemeinen viel zu umständlich). Hier hilft uns das Palatogramm. Mit ihm können wir die Unterschiede der Muskelspannung demonstrieren: bei der schlafferen Muskelspannung für ungespanntes [d] haben wir eine größere Berührungsfläche und bei der stärkeren Muskelspannung für das gespannte [t] haben wir eine dementsprechend

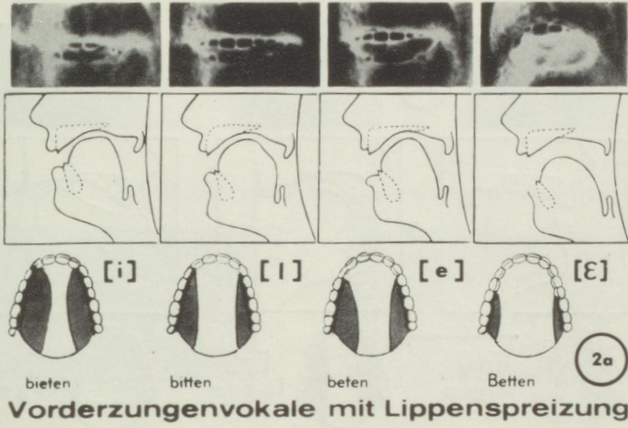


Abb. 1.

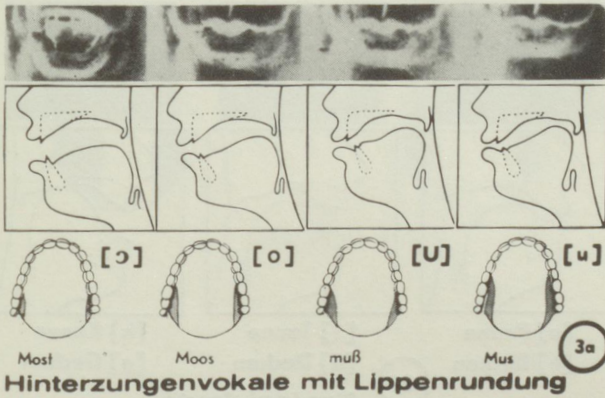


Abb. 2.

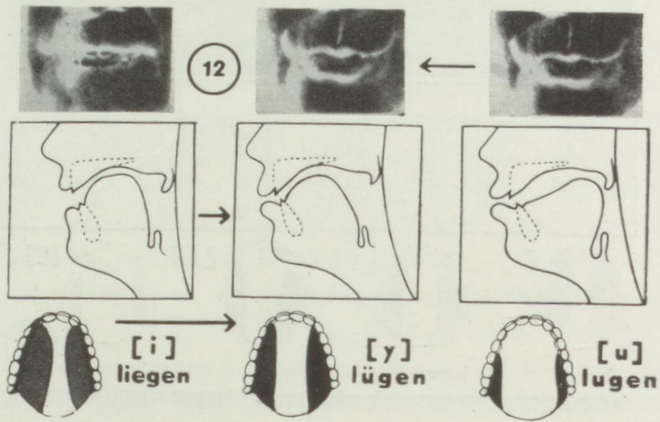


Abb. 3.

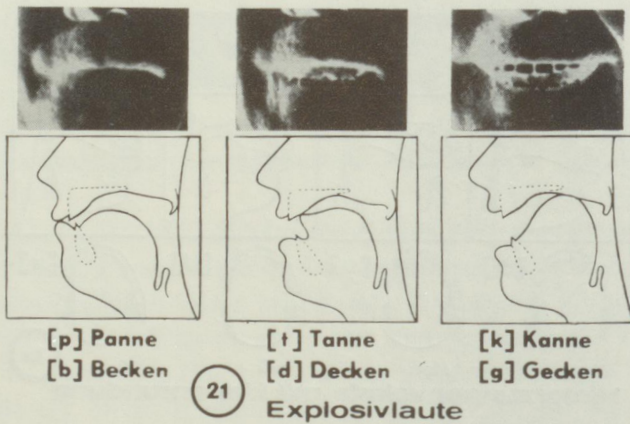


Abb. 5.

Martens: Optische Unterstützung akustischer Sprachsignale

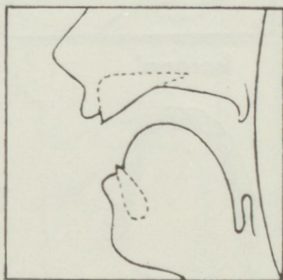
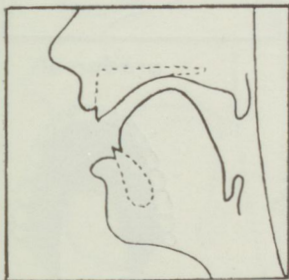
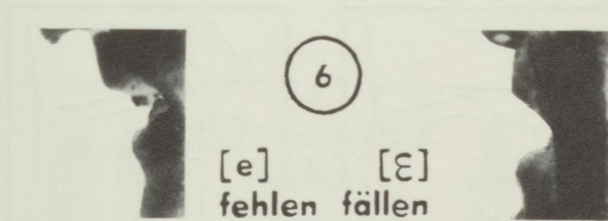
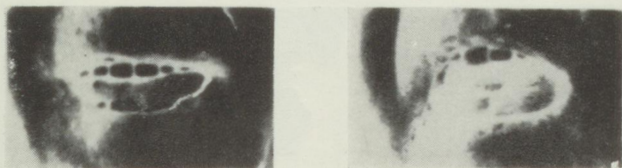
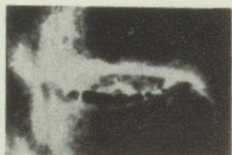
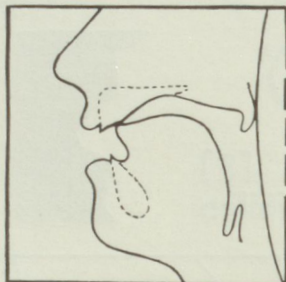


Abb. 4.

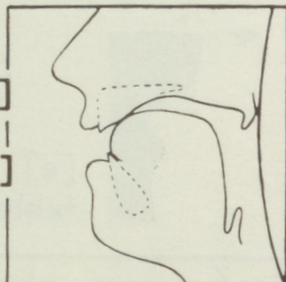


22



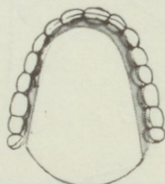
[t]

[d]

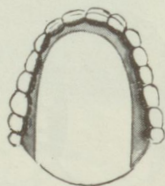


koronal

prä dorsal



[t]
Teich



[d]
Deich

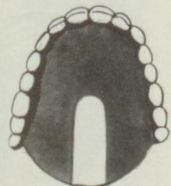
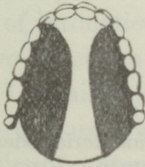
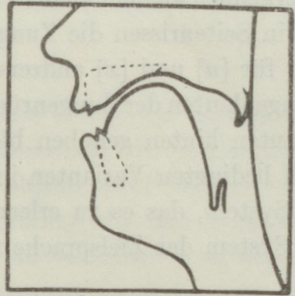
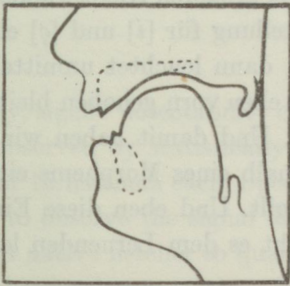


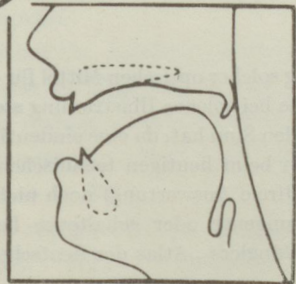
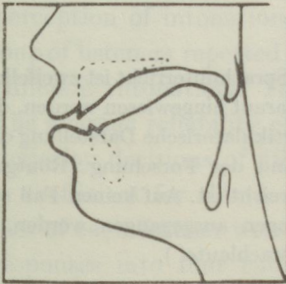
Abb. 6.



[i] → [ç]
riechen



29a



[u] → [x]
suchen

Abb. 7.

kleinere Berührungsfläche. Beides kann der Lernende nach der Demonstration und Erläuterung des Palatogramms meist sehr schnell nachempfinden. Damit haben wir ein weiteres Oppositionspaar der distinktiven Faktoren im phonologischen System der deutschen Sprache verdeutlicht: „gespannt“ gegen „ungespannt“. Dies hat noch dazu den Vorteil, daß wir uns lösen können von der oft irreführenden Verabsolutierung des Oppositionsbegriffs „stimmlos“ — „stimmhaft“.

Abb. 7: Die Unterschiede zwischen [ç] und [x] lassen sich zum Teil natürlich schon von außen her erkennen an dem kleineren oder größeren Kieferwinkel. Sieht man aber nebeneinander in Seitenrissen die Zungenstellung für [t] und [ç] einerseits und die Zungenstellung für [u] und [x] andererseits, dann leuchtet unmittelbar ein, warum nach Vorderzungenlauten der Zungenrücken eben vorn gehoben bleibt, dagegen nach Hinterzungenlauten hinten gehoben bleibt. Und damit haben wir wiederum diese kombinatorisch bedingten Varianten innerhalb eines Morphems eingeordnet in das phonologische System, das es zu erlernen gilt. Und eben diese Einordnung in das phonologische System der Zielsprache macht es dem Lernenden leichter.

DISCUSSION

Dvončová:

Ich habe die Frage, ob Herr Professor auch mit dem Filmmaterial oder nur mit Bildern bei dem Unterricht arbeitet. Wie ist seine Meinung über die Arbeit mit dem Filmmaterial.

Ezawa:

Die Bedeutung solcher optischen Mittel für den Sprachunterricht ist zweifellos nicht zu leugnen. Andererseits sollte bei solcher Illustrierung stets darauf hingewiesen werden, daß sie insofern nur einen orientierenden Sinn hat, da eine eindeutige artikulatorische Darstellung der Organstellungen von Sprachlauten beim heutigen technischen Stand der Forschung (Röntgenfilmaufnahme und ihre einwandfreie Auswertung) noch nicht erreicht ist. Auf keinen Fall sollte jedenfalls von der Analyse gesungener oder gehaltener Lautungen ausgegangen werden, wie es manchmal geschieht (vgl. Wänglers „Atlas der deutschen Sprachlaute“).

Martens:

Ad Dvončová: Zur Demonstration von Bewegungen ist der Film nicht durch Diapositive zu ersetzen. Für pädagogische Zwecke ist es jedoch oft nützlicher, ein typisches, sich nicht bewegendes Bild zu zeigen, weil es sich besser einprägt. Im übrigen sind solche bildlichen Darstellungen des Lautungsvorgangs lediglich didaktische Maßnahmen, die dem Lernenden helfen sollen zu begreifen, was er normalerweise nicht sehen und auch nur schwer ertasten kann. — Pädagogisch hat das Bild — und somit eben auch das Diapositiv — den Vorteil, daß es so lange gezeigt werden kann, wie der Lehrende es in der jeweiligen Situation für erforderlich hält.

Ad Ezawa:

1. p/b und t/d und k/g müssen wir im Deutschen auch dann als Oppositionspaare betrachten, wenn wir nur eine kleine Anzahl von Minimal-Paaren finden können.

2. Bei Vokalen läßt sich die „mittlere Phase“ für pädagogische Zwecke auf jeden Fall fruchtbar einsetzen, auch wenn wir uns auf den Standpunkt stellen, wir wüßten nicht genau was für die deutschen Vokale typisch sei.

QUANTIFICATION OF SIGNALS IN THE INTONATION OF AMERICAN ENGLISH BY HUMAN LISTENERS

ALESSANDRA MASSA*—DIANNE M. STEVENS**—YUKIO TAKEFUTA***

In this study, signal¹ detectability of native listeners in the intonation of American English was observed by statistically summarizing the psychophysical responses of three groups of 16 listeners each. Specific objectives of the present experiment were threefold: (1) to observe the signal detectability of a native listener, (2) to observe the ability of a native listener to quantify an intonational signal, and (3) to test the feasibility of the simple summation of component signal strengths to obtain the total signal strength of intonation in English sentences.

Twenty native American speakers and three groups of 16 native American listeners participated in this experiment. Each speaker recorded forty English sentences. Each sentence was paired with the same sentence spoken by a different speaker to make up 400 pairs of sentences. These 400 stimuli were presented to three groups of listeners for perception of intonational signals.

The first group of listeners reported whether the paired sentences were spoken with the same or different intonations. The listeners disagreed often in this apparently simple task. The average reliability of the responses of one listener was .62. Mean responses or majority decisions of 16 listeners, however, were sufficiently reliable. The reliability coefficient for this decision was .96.

The responses of each listener were then compared with the majority decisions to classify his responses into four categories: "HIT", "CORRECT REJECTION", "MISS", and "FALSE ALARM". Analysis of the responses of all listeners in the first group is presented in Table 1. The index of signal detectability² varied from 1.6 to 4.2, with an average of 2.4.

The second group of listeners reported the apparent magnitude of signal strength

* Università Cattolica del Sacro Cuore, Milano, Italy

** Denison University, Granville, Ohio, U.S.A.

*** The Ohio State University, Columbus, Ohio, U.S.A.

¹ A signal in intonation was defined in this study as any perceivable difference in the intonations of paired sentences.

² The index of detectability is a measure of the observer's sensory capabilities in a signal detection experiment and is obtained from the information of the HIT RATE and the FALSE ALARM RATE with reference to the table of cumulative normal distribution.

QUANTIFICATION OF SIGNALS IN THE INTONATION OF AMERICAN ENGLISH BY HUMAN LISTENERS

ALESSANDRA MASSA*—DIANNE M. STEVENS**—YUKIO TAKEFUTA***

In this study, signal¹ detectability of native listeners in the intonation of American English was observed by statistically summarizing the psychophysical responses of three groups of 16 listeners each. Specific objectives of the present experiment were threefold: (1) to observe the signal detectability of a native listener, (2) to observe the ability of a native listener to quantify an intonational signal, and (3) to test the feasibility of the simple summation of component signal strengths to obtain the total signal strength of intonation in English sentences.

Twenty native American speakers and three groups of 16 native American listeners participated in this experiment. Each speaker recorded forty English sentences. Each sentence was paired with the same sentence spoken by a different speaker to make up 400 pairs of sentences. These 400 stimuli were presented to three groups of listeners for perception of intonational signals.

The first group of listeners reported whether the paired sentences were spoken with the same or different intonations. The listeners disagreed often in this apparently simple task. The average reliability of the responses of one listener was .62. Mean responses or majority decisions of 16 listeners, however, were sufficiently reliable. The reliability coefficient for this decision was .96.

The responses of each listener were then compared with the majority decisions to classify his responses into four categories: "HIT", "CORRECT REJECTION", "MISS", and "FALSE ALARM". Analysis of the responses of all listeners in the first group is presented in Table 1. The index of signal detectability² varied from 1.6 to 4.2, with an average of 2.4.

The second group of listeners reported the apparent magnitude of signal strength

* Università Cattolica del Sacro Cuore, Milano, Italy

** Denison University, Granville, Ohio, U.S.A.

*** The Ohio State University, Columbus, Ohio, U.S.A.

¹ A signal in intonation was defined in this study as any perceivable difference in the intonations of paired sentences.

² The index of detectability is a measure of the observer's sensory capabilities in a signal detection experiment and is obtained from the information of the HIT RATE and the FALSE ALARM RATE with reference to the table of cumulative normal distribution.

Table 1. The percentage of "HIT", "MISS", "CORRECT REJECTION", and "FALSE ALARM" in the responses of a signal detection experiment of native listeners in the intonation of American English. The index of signal detectability was obtained by referring to a published table of cumulative normal distribution.

Listeners	1	2	3	4	5	6	7	8
Hit	33.5	44.1	36.7	40.4	44.1	50.0	46.8	41.0
Miss	16.5	5.9	13.3	9.6	5.9	0	3.2	9.0
Correct Rejection	48.9	48.9	47.3	44.1	46.3	29.3	48.9	46.8
False Alarm	1.1	1.1	2.7	5.9	3.7	20.7	1.1	3.2
Hit Rate	67.0	88.2	73.4	80.8	88.2	100.0	93.6	82.0
F. A. Rate	2.2	2.2	5.4	11.8	7.4	41.4	2.2	6.4
Index of Signal Detectability	2.5	3.2	2.2	2.1	2.6	4.2	3.5	2.4

Listeners	9	10	11	12	13	14	15	16	Average
Hit	49.5	36.2	45.7	39.9	48.4	49.5	43.6	48.4	43.6
Miss	0.1	13.8	4.3	10.1	1.6	0.1	6.4	1.6	6.4
Correct Rejection	45.7	49.5	45.2	38.8	47.3	34.6	43.1	46.3	44.7
False Alarm	4.3	0.1	4.8	11.2	2.7	15.4	6.9	3.7	5.3
Hit Rate	99.0	72.4	91.4	79.8	96.8	99.0	87.2	96.8	87.2
F. A. Rate	8.6	0.2	9.6	22.4	5.4	30.8	13.8	7.4	10.6
Index of Signal Detectability	3.7	3.5	2.7	1.6	3.5	2.8	2.3	3.3	2.4

by assigning a number to each set of paired intonations by the method of equal-appearing intervals using a 9-point scale. The average reliability of the responses of one listener in this task was .47. The reliability of the mean responses of 16 listeners was .92. The measures of signal strength expressed by the mean responses of the listeners of this group were compared with measures³ obtained from the responses of the listeners of the first group. The two measures correlated highly with each other with the coefficient of .90.

The third group listened only to those pairs of intonations which were unanimously reported either as similar or different by the listeners of the first group. The sentences they heard, however, had been truncated into each of the constituent words by an electronic switch device and randomized in the order before being presented to the listeners. Thus the listeners heard only a portion of the entire intonation and re-

³ The number of listeners who reported that the pair of sentences were different in intonation was interpreted as the relative measure of signal strength for any pair of intonations.

ported the perceived magnitude of signal strength in these pairs. The listeners again used the 9-point scale of equal appearing intervals in reporting their responses. The average reliability of the responses of one listener in this task was .66. The reliability of mean responses of 16 listeners was .97. The mean response for each portion of the intonation was then reassembled in such a manner as to restore the entire pair of intonations. The average response of signal strength of all the constituent parts was compared with the mean response given by the listeners of the second group to the entire intonation. Twenty-eight pairs of intonations which were unanimously decided upon by the 16 listeners of the first group to be either different or similar (Figure 1, top) varied according to a measure of signal strength reported by listeners of the

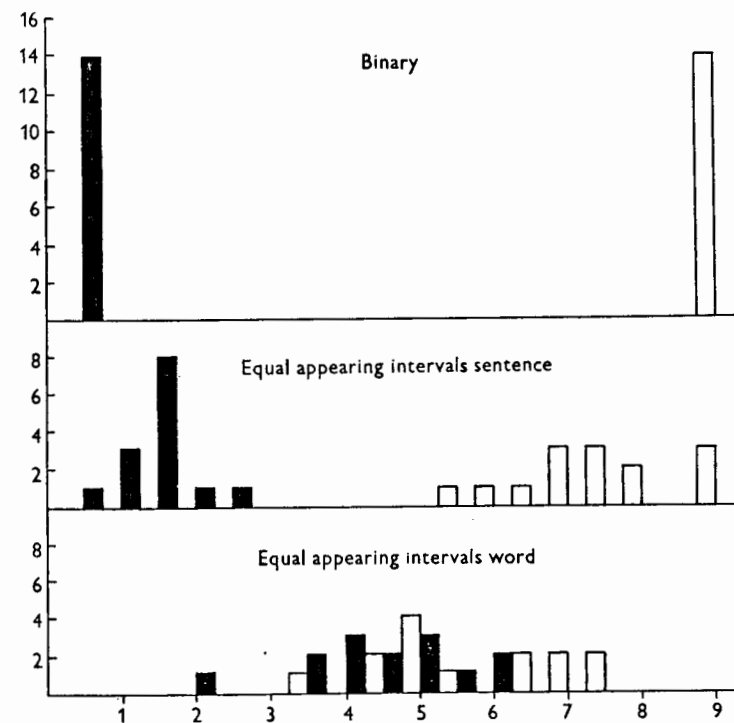


Fig. 1. Quantification of signals in the intonation of American English. The ordinate is the frequency and the abscissa is the 9-point scale of equal-appearing intervals. Twenty-eight sample pairs of intonations which were unanimously decided to be either different or similar (top) varied in the mean signal strengths estimated by the 9-point scale of equal-appearing intervals (middle). The average of the signal strengths of the constituent parts could not keep the dichotomy of the same samples (bottom).

second group. Dichotomy, however, of those sample stimuli was kept intact (Figure 1, middle). A simple summation or average of the measures of signal strength for all of the constituent parts of the intonation almost completely mixed the dichotomous samples (Figure 1, bottom).

CONCLUSION

A native listener, as is shown by the coefficient of reliability and the indices of the signal detectability, was not very competent either in detection of the signals or in quantification of signal strength in the intonation of American English. Mean responses of 16 listeners, however, could reliably detect or quantify the signal. The signal strength of the entire intonation was found not to be a simple summation or the average of the signal strengths of the constituent parts of the intonation as perceived independently from other parts.

REFERENCES

- Lieberman, P. "On the Acoustic Basis of the Perception of Intonation by Linguists," *Word*, 21, 1965, 40—54.
- Schwartz, J. R., and Takefuta, Y. "Detection of Accentual Signals by American and Japanese Listeners," Paper prepared for the 43rd Annual Convention of the American Speech and Hearing Association, Chicago, Illinois, U.S.A., 1967.
- Swets, J. A. (ed.). *Signal Detection and Recognition by Human Observers*. New York: John Wiley & Sons, 1964.
- Takefuta, Y. "A Study of Relative Efficiency of Acoustic Parameters in the Intonational Signal of American English." Unpublished Ph. D. dissertation. The Ohio State University, Columbus, Ohio, U.S.A., 1966.

ACKNOWLEDGMENT

This research was supported by a grant from the National Science Foundation (NSF GY-842, Undergraduate Research Participation Program) to The Ohio State University Department of Speech.

CZECH QUANTITY IN PARADIGMATIC AND SYNTAGMATIC PROCEDURES

LADISLAV MATĚJKA*

1. The lexicon and the grammar in the majority of Czech dialects employ a vowel system habitually described in terms of quantitative correlation. Native Czechs normally do not confuse the name of the letter $\acute{a}/a_1/$ ¹ or $\acute{u}/u_1/$ with the conjunction $a/a_2/$ or the preposition $u/u_2/$, or $d\acute{a}l$, meaning 'distance' with dal , meaning 'he gave'. or $p\acute{a}lívá$ (burns) with $paliva$ (of the fuel) and $p\acute{a}lívá$ (burning). Since quantity is distinctive in accented and unaccented syllables, it plays a part in the differentiation of monosyllabic as well as polysyllabic words. In this way, for example, $/mi_2li_2/$ (they washed) is distinguishable from $/mi_1li_1/$ (he deceives) which in turn is distinguishable from $/mi_2li_1/$ (nice) and $/mi_1li_2/$ (of the mile). The number of possible combinations grows with the number of syllables, although, as Trnka (1966)² has pointed out, not all possible combinations are equally utilized. It is noteworthy that in Czech poetic usage quantitative distinctions do not prevent acceptable rhyming; a rhyme $zahrad\acute{a}$ — $zahrát$ is even cited by the Czech aesthetician Mukařovský as an example of complete homophony („slova úplně stejnozvúčná“).³

2. The literature on Czech vowels is quite extensive. The phonologically oriented inquiries of the early days of linguistic functionalism were naturally concerned with the status of the quantity within the sound system. Mathesius (1929)⁴ was in favor of interpreting the quantity of Czech vowels as a modifying feature (“un élément modificateur”) which changes the functional aspect of a phoneme without changing it into a different phoneme (“qui ne fait pas passer celui-ci à un phonème différent”). On the other hand, Jakobson 1931/62)⁵ insisted on the distinction between short and

* University of Michigan.

¹ Subscripts 1 and 2 denote long and short vowels respectively. Cf. Jakobson R., and Halle M. “Tenseness and Laxness”, *R. J. Selected Writings*, I (1962) as well as Fant, G. “Theory of Distinctive Features”, *STL-QPSR-4/1966*.

² Trnka, B., “The Distribution of Vowel Length and its Frequency in Czech”. *Prague Studies in Mathematical Linguistics*, 1, (1965).

³ Mukařovský J., *Kapitoly z české poetiky*, 2 (Prague, 1948), 40.

⁴ Mathesius V., „La structure phonologique du lexique du tchèque moderne.“ *Travaux du Cercle Linguistique de Prague*, I (1929), 71.

⁵ Jakobson, R., “Phonemic Notes on Standard Slovak”, in *Selected Writings*; I (s-Gravenhage. 1962), 224. Cf. „Z fonologie spisovné slovenštiny“, *Slovenská miscellanea* (Bratislava, 1931).

long phonemes and included into his long series also the diphthong "ou" which he characterized as "an indecomposable phoneme". In accordance with his concept of structure, he put a particular emphasis on his argument that "it is the diphthong as a whole which is endowed with the quantitative feature of length and not its components *per se*."

3. Phonetically oriented studies, on the other hand, were primarily concerned with the articulatory characteristics of Czech vowels. Frinta⁶ in 1909 pointed out that short and long vowels in Czech are differentiated not only by the degree of duration but also by the degree of openness (or narrowness). In 1928, Chlumský⁷ submitted the length of Czech vowels to accurate measurements and in detailed statistics disclosed the relativity of duration in all vowels whether long or short. He also took into account various attendant conditions such as word stress, emphatic stress, tempo of the utterance and certain aspects of personal speech habits. Moreover, he undertook to study the role of various consonants and consonantal groups and demonstrated that a vowel which is identified as being long in certain syllabic structures can have, in fact, shorter duration than a vowel which is identified in a different environment as being short. The long *á* in the sentence *Jde k Pátečkovům* was found, for example, to be two centiseconds shorter than the short *a* in the phrase *slovo ba*. Hence the existence of short-long (L_s) and long-short (S_l) vowels in Czech was given an exact documentation.

4. Although numerous inquiries concerning Czech vowels have illuminated many essential problems, the challenge has not yet diminished. Some of the most crucial questions pertain, in fact, to the relativity of duration and to its bearing on the distinctive values of the Czech vowels. In the inspiring *Preliminaries to Speech Analysis* (1955)⁸ a suggestion was made that "in the Czech *pravá práva* 'true rights', the first vowel of the first word is identified as short in relation to the second, long vowel, while the second word displays the inverse relation." In other words, to use the reformulation in *Fundamentals of Language* (1956)⁹, "the quantity of a vowel may be established only in relation to the quantity of other vowels within the context or to the subsequent consonants."

While this stimulating observation about the syntagmatic, context-oriented procedures opens new vistas, it does not explain how *práva* (he/she/it washes) is distinguishable from *prava* (right), or *pálivá* from *paliva*; nor does it explain why the three *a*'s in *chvátává* are identified as long while the three *a*'s in *Svatava* are identified as short. Moreover, the syntagmatic, context-oriented comparison cannot explain the fact that a native speaker of Czech can produce and identify short and long vowels

in isolation and can use the very sound of the vowels as their names by calling them *a, á, e, é, i, í, o, ó, u, ú*. The adequate identification is obtained even if the vowel is carefully isolated from other utterances of the same speaker or to a considerable degree lengthened or shortened on a mechanical stretcher by changes in speed.

5. That a Czech native speaker is able to distinguish isolated short and long vowels in the absence of their long counterparts in the context, discloses the relevance of paradigmatic procedures oriented towards the system *in absentia*. If the explanatory statement is to be focused on the perceptual aspects, the question arises, whether such a paradigmatic identification of quantity with the help of a system *in absentia* primarily depends on the perception of duration or rather on some inherent features such as tenseness vs. laxness with relative duration as a concomitant factor. If the quantitative value of a vowel were primarily determined within the system by its duration, a class of shortness would be distinguished from the class of length regardless of the intensive variability (and overlapping) of their actual manifestation in the context. On the other hand, if the decisive information would primarily depend on the qualitative features, Czech vowel quantity should be more properly called vowel quality.

There are, however, good reasons to assume that Czech vowel quantity is determined within the system neither by any inherent nor by any prosodic feature *per se*, but by a structural value identifiable by an interaction of both syntagmatic and paradigmatic procedures. Such an interaction is always to be expected whenever variable duration and other factors, involved in the articulatory manifestation, interfere with the paradigmatic procedures. Since the quantity of a Czech vowel cannot be established only in relation to other vowels or consonants in the context, an interaction of syntagmatic and paradigmatic procedures seems to enable coordination of the contextual data with the data which can be derived only by means of a system *in absentia*, that is to say, in the memory.

DISCUSSION

Frinta:

Les voyelles courtes en tchèque peuvent être allongées par l'emphase, mais les longues ne se raccourcissent pas. La perception de la longueur chez les étrangers dépend de la circonstance de leur langue à eux. Si elle possède aussi la longueur phonologique, elle est perçue et bien imitée. Au contraire p.ex. les Polonais distinguent difficilement la longueur en tchèque. Les Russes prolongent les voyelles accentuées, mais pas les atones en parlant tchèque.

⁶ Frinta A., *Novočeská výslovnost* (Prague, 1909).

⁷ Chlumský J., *Česká kvantita, melodie a přízvuk* (Prague, 1928).

⁸ Jakobson, R., Halle, M. and Fant, G., *Preliminaries to Speech Analysis* (Cambridge, 1955)²

14.
⁹ Jakobson R., and Halle M., *Fundamentals of Language* (The Hague, 1956), 25.

AUDITORY PERCEPTUAL SKILLS OF CULTURALLY DISADVANTAGED CHILDREN

FREEMAN McCONNELL* AND JOE B. ROBERTSON**

In the pre-verbal stages of language learning, receptive language results from the process of relating incoming auditory symbols with appropriate visual, kinesthetic and other sensory images. The internalization of such symbols only becomes possible when meaningfulness is linked with them. This process of association presupposes that the child has perceptual abilities which make him capable of attending to, discriminating, recognizing and retaining sensory input.

Recent investigations (Deutsch, M., 1963; Weaver, 1963; Deutsch, C., 1964; Clark and Richards, 1966; and McConnell and Horton, 1966) have suggested that culturally deprived children have difficulty in auditory perceptual skills and are below age level in comprehending auditorially. The hypothesis seems tenable that environmental conditions could preclude the acquisition of meaningful auditory experience in the preschool years, thereby altering the development of basic listening skills. 'Tuned out' auditorially, such a child might demonstrate insensitivity to verbal input, and hence encounter much difficulty in auditory symbolic learning when placed in a school environment where auditory experience is crucial. The fact that the study of auditory-perceptual functioning in such children in the United States has been mainly limited to the investigation of speech sound discrimination suggested a need for further study involving other types of auditory skills. The present study was, therefore, planned to compare a group of culturally disadvantaged preschool children with two groups of culturally advantaged, or middle-class, children on tasks of auditory word memory and phonemic synthesis, as well as speech discrimination. In addition to delineating differences between middle class and lower class children on these specific types of auditory-perceptual skills, it was planned to determine if any differences observed might be related primarily to social class factors or to ethnic differences involved.

* The Bill Wilkerson Hearing and Speech Center, Vanderbilt University School of Medicine, Nashville, Tennessee, U.S.A.

** Hearing and Speech Center, Children's Hospital, Washington, D.C., U.S.A.

PROCEDURE

Three groups of 15 children each served as subjects for the study. The experimental group was composed of 15 culturally disadvantaged Negro children, so classified on the basis of the particular day care center in which they were enrolled and the neighborhood in which they lived. These subjects resided in two of the most economically depressed areas of the city, characterized by high rates of family mobility, juvenile delinquency, and school dropouts.

The two middle class groups, one Negro and one Caucasian, came from two private church-sponsored kindergarten programs and a state university demonstration-type kindergarten. Each of these children had at least one parent who was a college graduate, and parental occupations were mainly professional and business.

Each child was between four and six years of age. Hearing levels, screened in the speech frequency range (500—2000 cps) by pure tone audiometry, were required to be 20 dB or better (1964 ISO). Speech behavior and emotional and intellectual development were within the normal range on the basis of either psychological evaluations available or teacher evaluation.

A battery of three tests was given to assess auditory-perceptual abilities. The first of these was the Auditory Word Memory Test, modeled after the Visual Memory Test included at the four-year level in the Stanford-Binet Intelligence Scale, Form L-M. The subject was required to listen and retain a series of words in sequence, then repeated with one of the words omitted. The child had then to recall the auditory image of the missing word from the first set and repeat it to the examiner. The items were arranged in terms of increasing difficulty, ranging from two to five words per sequence from beginning to end.

The second test, Auditory Blending, was a subtest from the Marion Monroe Reading Aptitude Tests (1935). This 12-item test required the synthesis of a word when its individual phonemes were presented in correct order but with a time delay between each phoneme, as for example, *c-a-t*. The child was required to designate from a possible choice of three pictures the one representing the word the examiner had presented orally in a phonemically disconnected manner.

The third test administered was a speech sound discrimination test in four subtests, each of which contained 28 paired word stimuli which were either identical, or different in only one phoneme unit. The first word of the pair was presented to the left ear, while the second half of the pair was received in the right ear. The listening task was made increasingly difficult by shortening the time span between words until the two words of the pair were completely overlapping, and thus heard simultaneously in the two ears. The child's task was to point to the illustration representing the stimulus pair he had just heard. The test material was recorded on tape at a standard recording level and presented through an Ampex stereophonic tape recorder from the control room of a two-room sound isolated test suite. The output of each of the two channels was fed into an amplifying system, calibrated to deliver the recorded message at 70 dB SPL at the earphones.

RESULTS

Analysis of the data from the foregoing tests, which is summarized in Table 1, showed that the culturally deprived group were markedly poorer on the three auditory tasks compared to each of the two middle class groups. Although the two middle class groups did not differ significantly from one another on two of the tests, the Negro group fell below the Caucasian group on the Auditory Blending Test, for which

Table 1. Comparison of a culturally disadvantaged group of preschool children with two culturally advantaged groups on three tests of auditory abilities

	Word Mem. Mn (and S. D.)	Aud. Blend. Mn (and S. D.)	Sp. Snd. Discrim. Mn (and S. D.)
Culturally Disadvantaged—Negro (N-15)	4.8 (3.7)	4.6 (1.7)	65.5 (11.9)
Culturally Advantaged—White (N-15)	8.7* (2.4)	7.4 (1.8)	78.7* (10.2)
Culturally Advantaged—Negro (N-15)	8.0* (2.4)	6.0 (1.8)	75.7* (13.7)

* All observed differences between groups with the exception of cult. adv. white and cult. adv. Negro on Word Mem. and Sp. Snd. Discrim. were statistically significant at .05.

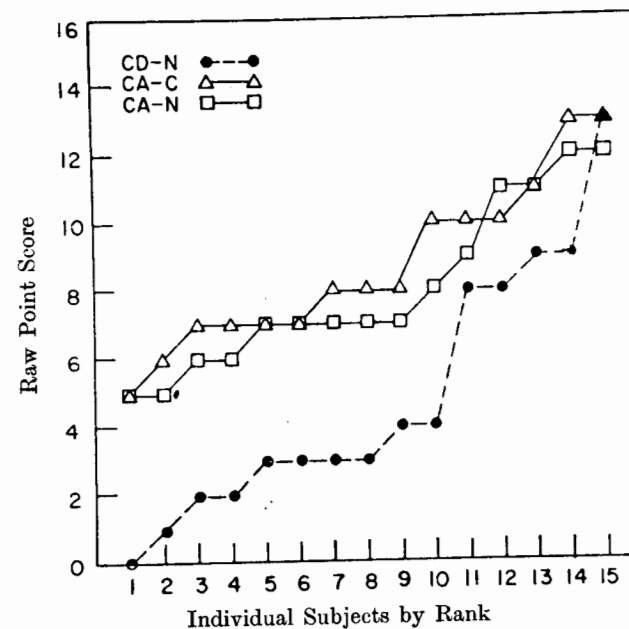


Fig. 1. Point scores obtained by each of the individual subjects on the test of Auditory Word Memory and plotted by rank from lowest to highest scores in each group.

their mean score was exactly halfway between the disadvantaged Negro group and the advantaged Caucasian group. Figures 1, 2, and 3 show the performance curves for the 15 subjects in each group, progressing from the lowest score plotted at the left to the highest score at the far right. The superiority of the two culturally advantaged groups compared to the one culturally disadvantaged group is quite apparent. The reason for the difference between the Negro and Caucasian advantaged groups on

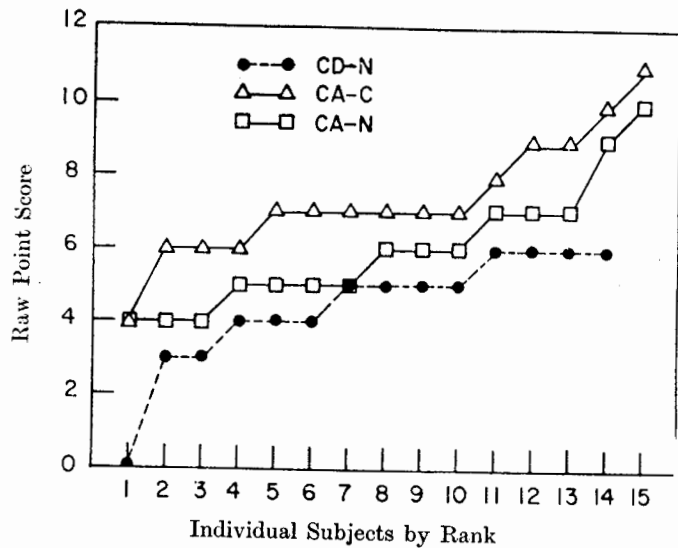


Fig. 2. Point scores obtained by each of the individual subjects on the test of Auditory Blending and plotted by rank from lowest to highest scores in each group.

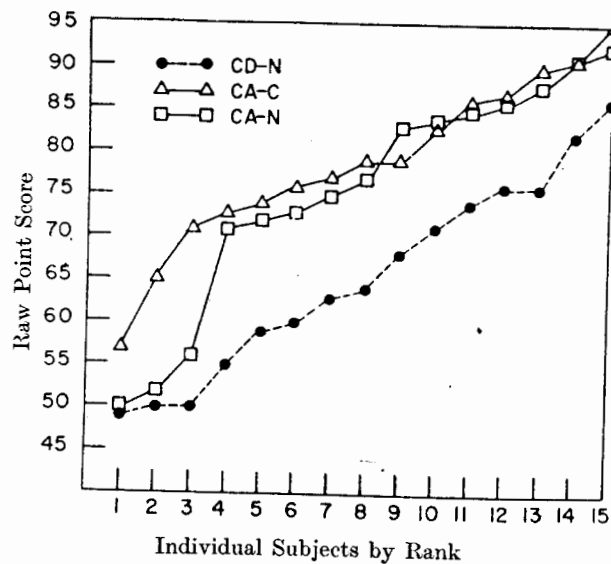


Fig. 3. Point scores obtained by each of the individual subjects on the test of Total Speech Sound Discrimination and plotted by rank from lowest to highest scores in each group.

Auditory Blending is not clear. It could be related either to Negro dialect factors or to other variables not controlled in this particular study. The study has suggested, however, that the environmental factor is of major significance in the development of auditory listening skills in the culturally deprived child.

SUMMARY AND DISCUSSION

In summary, the present study has confirmed findings reported by others that an early environment which is meager in much of the stimulus content surrounding the average middle-class American child may indeed result in poorly developed auditory skills. One may conjecture that language and vocabulary functioning will be reduced when auditory functioning is reduced, but what factors might be operating to produce such an auditory lag when hearing itself is normal? It is quite probable that most of these children come from homes in which the sound environment is extremely random and unstructured. For example, if a youngster has been living in a three-room home in a family of seven or eight people, the noise level in the home would very likely be quite high, and would consist of a barrage of extraneous sounds which the child would gradually learn to ignore. He thus blocks out auditory patterns from a very early age because they are so profuse he cannot handle them selectively at this stage in his auditory development. This blocking of auditory stimulation is thence a protection device for the young organism, and if persisting, may directly affect the child's capacity to transmit sensory stimuli by ear — the chief method by which every child learns language.

A problem yet unsolved is whether the basic auditory perceptual skills learned by most children in the first years of life, if unlearned because of the particular environmental milieu, would directly contribute to the child's lack of facility in language and poor educational achievement. The relation is indeed a logical one; at the present moment, we can only observe that the two deficiencies do exist in the same children.

REFERENCES

- Clark, Ann and Richards, Charlotte, Auditory discrimination among economically disadvantaged and nondisadvantaged preschool children. *Exceptional Children*, 33 (1966), 259—262.
- Deutsch, C., Auditory discrimination and learning: social factors. *Merrill-Palmer Quarterly*, 4 (1964), 277—296.
- Deutsch, M., The disadvantaged child and the learning process. In Passow, A. H. (editor), *Education in Depressed Areas*, New York: Teachers College, Columbia University (1963).
- McConnell, F., and Horton, Kathryn, Language characteristics of the culturally deprived child. In *Actas del 1er Congreso Americano de Hipoacusia y Lenguaaje*. Buenos Aires, Argentina: Foundation for the Help of the Deaf Child (1966), 584—589.
- Weaver, S., Interim Report: Psycholinguistic abilities of culturally deprived children. Mental Health Project grant 5-R11 MH 00765-04. Unpublished paper (1964).

SOME PROBLEMS OF OVER-ALL PATTERNING

RAVEN I. McDAVID, JR.*

When this paper was proposed, it was anticipated that the long-advertised Chomsky-Halle *The Sound Pattern of English* would be available for discussion. It is not. The discussion will therefore concentrate on the best-known and most successful attempt to establish an over-all phonological pattern accommodating the contrasts of all dialects of English, that of Trager and Smith in their *Outline of English Structure* (Norman, Oklahoma, 1951). All criticisms are data-oriented; except where otherwise specified, the data may be found in the collections of the Linguistic Atlas of the United States and Canada, which scholars are welcome to examine at the University of Chicago, even as editorial work proceeds.

1. The Trager-Smith analysis sets up for English a system of nine vowel phonemes, symmetrically distributed:

i	ɨ	ɨ̄
e	ə	o
æ	a	ɔ

There are three semivowels /y, w, h/ and 21 other consonants /p, t, k, b, d, g, č, ĵ, m, n, ŋ, f, v, θ, ð, s, z, š, ž, r, l/, four levels of pitch, four degrees of stress, and five degrees of juncture (including "normal" or "close" transition, defined as the absence of juncture). Long vowels and diphthongs are analyzed as sequences of short vowel + semivowel, following earlier suggestions by Henry Sweet and Prince Trubetzkoy. This analysis has received distinguished exegesis, notably by A. A. Hill (1958) and W. Nelson Francis (1958), and has enjoyed widespread and successful use in a variety of teaching situations.

2. It was inevitable that there would be challenges to the claims of Trager and Smith that their system can accommodate all contrasts in all varieties of English. Without recapitulating all arguments in details, we may summarize the kinds of criticisms that have been made.

3. Little comparable data is provided (or is available elsewhere) on variations in the suprasegmentals (stress, pitch, juncture), whether or not these are to be considered

* University of Chicago.

phonemes. Yet many observers have noted informally that regional and social differences exist in details of intonation and stress and possibly in the system.

4. There is no provision for contrasts between clear and dark varieties of /l/ and /n/, though these contrasts exist in some Scots dialects and possibly in the United States.

5. As more dialects are examined, especially in the southeastern United States, new vowel and diphthong contrasts are discovered that cannot be accommodated in the Trager-Smith pattern; this is particularly true of length and centering glide, which are lumped together in this pattern as postvocalic actualizations of /h/. Even in the New York City area there may be a minimal contrast between *maw* "stomach" and *moa* "extinct flightless bird". Expansion to take care of all such contrasts leads ultimately to fairly minute phonetic recording, a travesty on the phonemic principle.

6. The pattern may represent structural differences that do not exist. From Pennsylvania south, the syllabic nucleus of *food*, *spoon* and the like (which is elsewhere treated as /uw/) is usually a high-central rounded long vowel or upgliding diphthong [u; uu ɹ]. Trager and Smith would interpret this as /iw/; however in these same areas the syllabic nucleus of *put* (Trager's /u/), may also be a high-central or a fronted high-back rounded vowel, so that without multiplying entities one may suggest simply a difference in the phonetic quality of /uw/.

7. The pattern may not represent structural differences that do exist. Few middle-class Chicagoans have a high-central /i/ in contrast with high-front /i/ and high-back /u/; in their speech, phonetic [ɛ] occurs only under weak stress, in complementary distribution with strong-stressed [ɪ]. Similarly, many dialects have phonetic [a] and [ɔ] and intermediate sounds, but only in complementary distribution. The contrasting structures of such dialects may be buried in the symmetry of the pattern.

8. On balance, of course, the Trager-Smith analysis is still very serviceable; in fact, it is probably as successful and comprehensive a design of this type as can be made. Its incompleteness is noted, chiefly to remind us to be wary of all other such designs, however ingeniously devised and elegantly presented. Such designs are convenient abstraction — fictions, if we will — to be ruthlessly tested against the stubborn facts of independently recorded data.

DISCUSSION

Vachek:

Prof. McDavid's criticism of Trager-Smith's overall pattern of American English vowels can be endorsed. It was led from the same methodological positions as is held by the authors of the pattern themselves. But the idea of the overall pattern can be attacked on a still broader, more general basis. As a matter of fact the idea of such a pattern appears to be linguistically quite unfounded, as only such elements may be patterned (i.e. attributed to the same phonemic pattern) as actually co-exist in the given idiom. An overall pattern can, of course have a sociolinguistic significance, but then one has to do with an inventory of items, not with a real system.

EINRICHTUNGEN ZUR MEHRFACHANALYSE AKUSTISCHER SIGNALE

DIETER MEHNERT*

Im folgenden soll über 3 Gerätegruppen berichtet werden, die im Institut für Phonetik und Kommunikationswissenschaft der Humboldt-Universität Berlin entwickelt und hergestellt wurden. Dabei handelt es sich

1. um eine Weiterentwicklung des von Winckel beschriebenen Repetierzusatzes zu einem vielseitig verwendbaren selbständigen Gerät,
2. um eine bausteinartige Mehrzweck- u. Experimentieranlage für den Einsatz in phonetischen Laboratorien und
3. um ein Tonhöhenaufzeichnungsgerät, an das zur Registrierung Wachspapier-schreiber, Licht- u. Tintenstrahlzillografen angeschlossen werden können.

Zur Unterscheidung und Beschreibung einzelner Sprachlaute ist u. a. eine Zerlegung der Zeitfunktion erforderlich. Neben der Aufzeichnung verschiedener Meßdaten zur weiteren Auswertung kommt der Beurteilung der Sprachlaute mit dem Ohr nicht unwesentliche Bedeutung zu. Genau ist das nur dann möglich, wenn ein Lautabschnitt genügend oft wiederholt werden kann. Eine Einrichtung ist dafür besonders geeignet, wenn sie es gestattet, den interessierenden Teil aus einem Sprachsignal herauszuisolieren und dann ständig zur Beurteilung zur Verfügung zu stellen. Mit Hilfe eines umlaufenden Wiedergabekopfes an einer stehenden Bandschleife ist das einfach zu realisieren.¹ Wenn jetzt innerhalb des Umlaufes von 0—360° verschiedene Segmente geschaltet werden können, oder die Bandschleife den Umfang einstellbar tangiert, können genügend kleine Abschnitte analysiert werden.

An beiden Seiten des Repetiereinsatzes sind 2 Gleitschienen mit Einstellskalen angebracht, auf denen sich 2 Bandumlenkrollen verschieben lassen. Somit ist es möglich, das Band, also kleine Abschnitte des zu untersuchenden Materials, an den Wiedergabekopf bzw. an die Umschaltstelle bei 180° zu legen. Durch die Bandabhebung ist eine bessere Konzentration auf den bestimmten Lautabschnitt aus fortlaufendem Text möglich und man spart Schneid- u. Klebearbeit. Die Möglichkeit, Zeitabschnitte, wenn auch nur mit einer begrenzten Genauigkeit, zu messen, wurde damit erreicht, daß das zu untersuchende Band über Rollen läuft, auf die eine Meßskala angebracht ist.

* Humboldt-Universität Berlin, Inst. f. Phonetik und Kommunikationswissenschaft.

¹ Winckel, Repetierzusatzgerät zum Magnetofon für Lautuntersuchung. *Z.f. Phonetik u. allg. Sprachwiss.* 1951, Heft 1/2.

An Hand des Übersichtsschaltbildes — Modulation — sollen die technischen Einzelheiten des gesamten Gerätes erläutert werden. Das Gerät ist mit 3 verschiedenen Eingängen ausgerüstet und zwar einmal niederohmig symmetrisch, wie es zum Anschluß von Studiobandmaschinen erforderlich ist, dann hochohmig unsymmetrisch, z. B. zum Anschluß von Heimgeräten und symmetrisch niederpegelig, zum Anschluß von Kondensator- oder dynamischen Mikrofonen. Durch Mikrofonanschluß auf dem Gerät können Beispielbänder gleich durch entsprechende Ansage gekennzeichnet werden.

An das Repetiergerät können zum qualitativen Abhören von Beispielen dynamische Kopfhörer oder eine Tonsäule angeschlossen werden. Ebenso ist ein Abhören über einen eingebauten Lautsprecher möglich. Über einen Schalter sind die drei Betriebsarten des Gerätes Aufnahme/Wiedergabebetrieb, Repetierbetrieb und Verstärkerbetrieb einzustellen. Über einen Stufenschalter ist ein regelbarer beschleunigter Vor- u. Rücklauf des Bandes zum Aufsuchen einer bestimmten Bandstelle möglich. Die wichtigsten Bedienungselemente sind elektrisch gegeneinander verriegelt, damit bei evtl. Fehlbedienung kein Bandriß oder größerer Defekt entstehen kann. Zusätzlich ist das Gerät mit einem 6 dB Ausgang zum Anschluß von anderen Analysegeräten, z. B. einem Frequenzspektrometer, ausgerüstet.

Für jede Analyse oder weitere Untersuchung des komplexen Sprachsignals, die vom Band her durchgeführt werden soll, ist eine qualitativ hochwertige Tonbandaufnahme Voraussetzung. Es stehen dazu hochwertige Mikrofone, hochwertige Tonbandmaschinen und hochwertige Registriereinrichtungen zur Verfügung, jedoch bereiten häufig bei einem Versuchsaufbau die Zwischenglieder bezüglich Störspannungsabstand, Ein- u. Ausgangsimpedanzen, unterschiedlichen Pegeln, Schwierigkeiten. Bei dem Entwurf der Experimentieranlage ist davon ausgegangen worden, einige Bausteine zu schaffen, die allen Anforderungen, die sich aus der Experimentierpraxis ergeben, zu genügen, und die alle diese Zwischenglieder, Verstärker, Impedanzwandler, Dämpfungsglieder, Regelglieder usw. enthalten.

Die Anlage besteht aus dem Reglerfeld und zwei Normgehäusen, in die wahlweise $1/4$ Kassettens-Modulationsverstärker, -Verteilungsverstärker oder -Leistungsverstärker, z. B. zum Betrieb von Schleifenzillographen, eingesetzt werden können. Sämtliche Ein- u. Ausgänge aller einzelnen Elemente sind auf Trennklinken geführt und verschiedene Bauelemente sind schon zu Gruppen zusammengefaßt, wobei eine so vorgegebene Schaltverbindung, in der eine Trennklinke liegt, durch Stecken eines Klinkensteckers aufgetrennt werden kann. Alle Trennklinken sind auf der Rückseite der einzelnen Bausteine angeordnet, die Verbindung erfolgt über geschirmte doppeladrigte Klinkensteckerschnüre.

Das Übersichtsschaltbild — Modulation — zeigt die technischen Besonderheiten. Die ganze Anlage ist niederohmig symmetrisch ausgelegt. Es stehen 8 Eingangsgleitungen, umschaltbar auf 4 Kanäle, zur Verfügung, deren Ausgänge ebenfalls auf Klinken enden. Die Regler 1—4 sind zu einer Gruppe zusammengefaßt und entsprechende Entkopplungsglieder sind eingebaut. Über Regler 6 und 7 sind hochpegelige

Quellen anschließbar, die, falls sie unsymmetrisch sind, über eingebaute Übertrager zu stecken sind. Im Klinkenfeld sind Parallel-Klinkengruppen angeordnet, die mit jedem Ausgang oder Eingang, der im Versuchsaufbau mehrmals gebraucht wird, verbunden werden können.

Im Reglerfeld sind Abhörpunkte, die an den wichtigsten Teilen der Schaltung liegen, fest über einen Abhorschalter mit einem Abhörverstärker und Kontrolllautsprecher verbunden. Durch Tastendruck können diese Punkte nacheinander abgehört werden. Zum schnellen Durchprüfen einer gesteckten Versuchsanordnung ist das sehr vorteilhaft.

Auf dem Reglerfeld ist ein Anschluß für ein Kommandomikrofon angebracht, ebenso ist das Gerät mit einer Lichtzeichenanlage für Mikrofone ausgestattet. Der Fremdspannungsabstand der Anlage liegt bei ≥ 60 dB und genügt somit den geforderten Werten.

Die Einsatzmöglichkeit der Anlage ist vielseitig und reicht von einfachen Mikrofonaufnahmen über Mehrkanalaufnahmen bis zum Aufbau umfangreicher Meßanordnungen, u. a. sind damit eine Anlage für Vokalsynthese und eine Anlage für Filteroszillografie aufgebaut worden.

Ein wesentliches Merkmal der Sprache ist der Verlauf der Tonhöhe. Wenn statistisches Material in größerem Umfange bearbeitet werden soll, kann ohne direkte Aufzeichnungsverfahren nur wenig ökonomisch gearbeitet werden. Die Tonhöhenaufzeichnung in Verbindung mit Schalldruck, Lautstärke und Zeitmarke, hat sich als geeignetes Hilfsmittel bei phonetischen Untersuchungen erwiesen.

Von besonderem Vorteil ist die Aufzeichnungsart der Meßgrößen, die sofort ausgewertet werden kann, das sind die Wachspapier- und die Tintenstrahlschreibung. Für den Anschluß dieser beiden Registriergeräte und darüber hinaus für den Anschluß von Lichtstrahloszillografen ist das Tonhöhenaufzeichnungsgerät ausgelegt.

Neben der Einrichtung zur Ermittlung der Sprachgrundfrequenz² sind ein Frequenzmesser, ein lg-Verstärker zur Aufzeichnung des Pegels, ein elektronischer Gleichspannungswandler zum Anschluß des Brüel- und Kjaer-Schreibers und die entsprechenden elektronisch stabilisierten Netzgeräte eingebaut.

Des weiteren werden ein zweiter lg-Verstärker mit Ohrkurvenfilter zur Aufzeichnung des frequenzbewerteten Schalldruckes, ein umschaltbarer Zeitmarkengenerator, damit jeweils die optimale Zeitmarke für den gewählten Vorschub des Registriermaterials geschrieben werden kann, und ein Eich- und Einstellgenerator für das gesamte Gerät, eingesetzt.

Mit Hilfe des letzteren soll vor jeder Messung eine treppenförmige Frequenz- und Pegelkurve geschrieben werden können, die die Maßeinheiten für die Ordinate liefert und zur Auswertung des untersuchten Materials notwendig ist.

² Tscheschner/Pohlink, Ein Beitrag zur Messung der Stimmbandgrundfrequenz, *Wiss. Z. TU Dresden* 14, 1965, Nr. 6.

Mehnert: Einrichtungen zur Mehrfachanalyse akustischer Signale

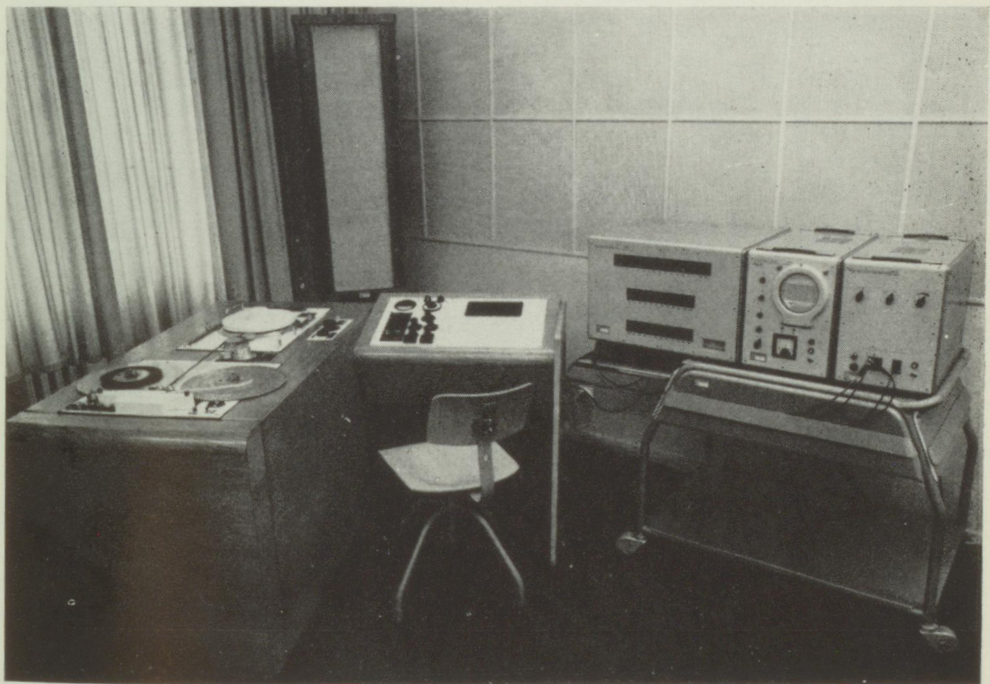


Bild 1: Repetiergerät in Verbindung mit Frequenzspektrometer.

Mehnert: Einrichtungen zur Mehrfachanalyse akustischer Signale

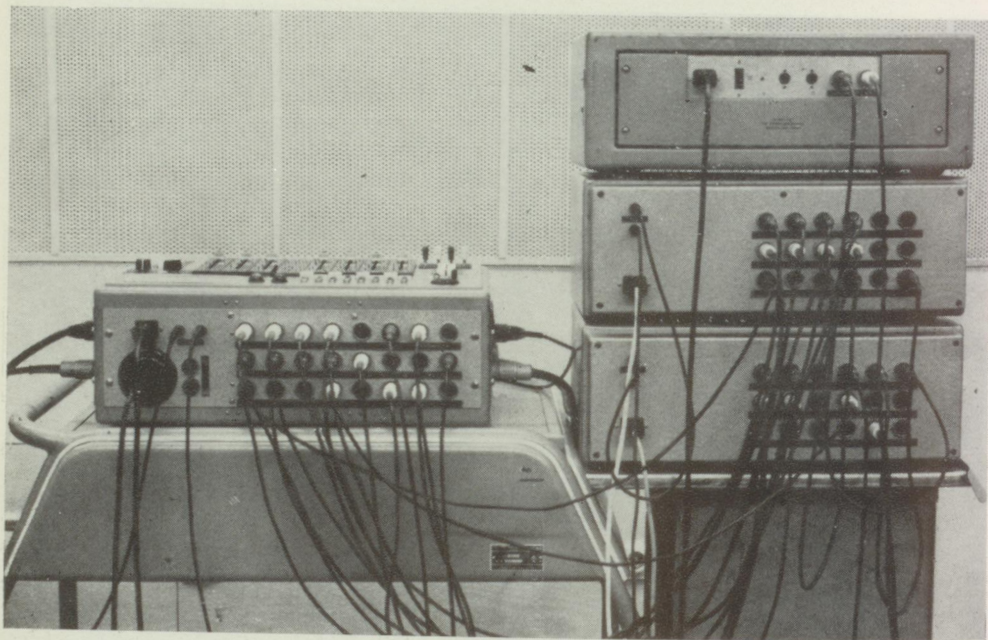


Bild 2: Bausteinförmige Experimentieranlage (Rückseite)

Mehnert: Einrichtungen zur Mehrfachanalyse akustischer Signale



Bild 3: Bausteinförmige Experimentieranlage — Reglerfeld

EINIGE KRITERIEN FÜR DIE ERKENNUNG AKUSTISCHER SIGNALE

GEORG F. MEIER—JOHANNES KYRITZ

Die automatische Zeichenerkennung wird derzeit sowohl in biologischen wie auch in technischen Systemen erforscht. Für die Erforschung des technischen Vorganges, z. B. für automatische Diktieranlagen, kann ein rein technischer Weg eingeschlagen werden, d. h. daß die akustischen Parameter von Silben bzw. Lautübergängen und ihre Verarbeitung durch analoge oder digitale Datenverarbeitungsanlagen untersucht werden. Ein anderer Weg, der zugleich den menschlichen Kommunikationsvorgang, besonders die Dekodierung erfassen kann, ist der *bionische Weg*. Besonders der letztgenannte Weg soll nachfolgend behandelt werden. In dieser kurzen Mitteilung sollen nur 3 Probleme und Ergebnisse herausgegriffen werden:

1. ROLLE DER SEMANTIK BEI DER SIGNALAUFNAHME

Über die Relevanz oder Redundanz einzelner Elemente (Segmente) eines Textes können einige allgemeine Gesetze aufgestellt werden:

1.1 Jeder Satz hat als Ganzes eine Kontaktbedeutung, die oft nur durch ein einziges Element (Fragepartikel) oder durch wenige seiner Elemente getragen wird (Inversion, Frageton usw.). Man spricht auch vom „kommunikativen Aspekt“.

1.2 Jeder Satz stellt einen Komplex von Bedeutungselementen (Noemen) dar, deren Träger die Lexeme, Lexemgruppen, Positionen der Lexeme, Toneme oder Tonemfolgen und andere Prosodeme sind.

1.3 Die Noeme bestimmter Syntagmen bilden untereinander Durchschnittsmengen aus den Vereinigungsmengen der am Syntagma beteiligten Sememe, wenn zwischen diesen Sememen Verträglichkeitsrelationen bestehen.

1.4 In der Ontogenese verknüpfen sich viele häufig vorkommende Lexeme zu dynamischen Stereotypen, so daß die Auslösung des Anfangsgliedes genügt, um den Rest des Stereotyps abfolgen zu lassen.

1.5 Die Satzgenerierung erfolgt vom Gedanken zur Form: bestimmte semantische Elemente oder Komplexe steuern die formalen Stereotypen. Lückenteste und Ratefolgeteste haben gezeigt, daß es semantische stochastische Prozesse (Markov-Ketten) gibt. Laute, Silben und teils auch Wörter können objektiv entfallen, ohne daß sie beim Perzeptionsprozeß vermißt werden und ohne daß es ein Mißverständnis gibt.

1.6 Wirkliche Oppositionen in der aktuellen Rede gibt es nur, wenn der gesamte Kontext völlig gleich ist und der Text sich nur in dem betreffenden Element unterscheidet.

1.7 Phonetische Stereotypen, denen semantische Folgen zugeordnet sind, sind teilweise redundant. Die Länge des Kontextes ist proportional zur Redundanz einzelner Elemente. Lautverwechslungen, Überschreiten der Invarianzgrenzen und Homophonie erschweren das Verständnis umgekehrt proportional zum Informationswert der Umgebungselemente.

1.8 Die Phonemabgrenzung beruht nicht zwingend auf semantischen Kriterien, sondern auf akustischen (statistischen) Gewohnheiten.

2. ROLLE BESTIMMTER ABHÄNGIGKEITEN ZWISCHEN FORM UND FUNKTION

Alle formalen, d. h. grammatischen und prosodischen Erscheinungen müssen auf ihre kommunikative Funktion hin geprüft werden. Als Beispiel sei der Komplex der Intonation herausgehoben: Die physikalische Abhängigkeit eines Tones sei dargestellt durch: $I = f$ (Formanten, Grundfrequenz, Zeit, Amplitude) =

$$f(p_1, p_2, p_3, p_4).$$

Ergeben mehrere Töne zusammen eine Tonfolge: $I_1 I_2 I_3 \dots I_n$, so entsteht eine Tonfolge-Relation $R_{seq}(I_1, I_2, I_3, \dots, I_n)$. Für den einfachsten Fall einer Tonfolge mit zwei Tonelementen ergibt sich folgendes Abhängigkeitsbild:

$$\begin{aligned} \{R_{seq}[I_1 = f(p_1, p_2, p_3, p_4), I_2 = f(p'_1, p'_2, p'_3, p'_4)] = \\ = f[R_i(S_1, \#, S_2) + R_{prop}(S, 2S) + R_{struct}(W, S_i) + \\ + R_{struct}(S_i, \text{Syntagma} > W) + \text{voix}(Sp)] = \\ = f\{R_{loc}(Sp, Hr) + R_{psych}(Sp, Hr) + \\ + R_{isom}[(S_1, S_2), (N_1, N_2, \dots, N_m)] + Z_{physiol}(Sp) + \text{Intentio}(Sp) \dots\}. \end{aligned}$$

Dabei bedeutet: R = Relation, t = Zeit, $prop$ = Proportion, S = Silbe bzw. Tonträger, $struct$ = Struktur, W = Lexem, $voix$ = Stimmlage, loc = lokal, $psych$ = psychisch interpersonell, Sp = Sprecher, Hr = Hörer, $isom$ = Zuordnung, N = Noem, Z = Zustand

Diese Formel besagt, welche Faktoren bei dem Zustandekommen der Intonation hauptsächlich mitwirken.

Eine wichtige Rolle spielt dabei und bei vielen anderen Abhängigkeiten der Zeitfaktor.

3. BESTIMMUNG DES ZEITFAKTORS

Alle Messungen, seien es technische oder bionische, müssen objektiv sein. Ich verweise auf zahlreiche Arbeiten mit Segmentiergeräten und sogenannten Zeitfilteranlagen (wie z. B. von Tscheschner entwickelt). Es geht dabei zunächst um die Ermittlung bestimmter perceptiver Merkmale von Lautsegmenten in Abhängigkeit von ihrer zeitlichen Dauer, um die Feststellung des Einflusses von Verdeckungseffekten zeitlich

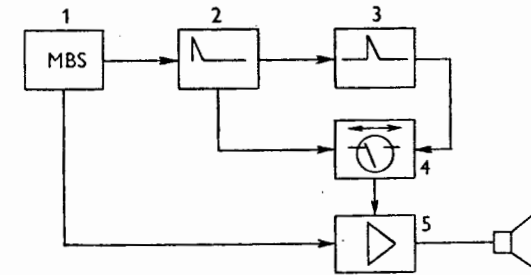


Fig. 1. Prinzipielles Blockschaftbild eines Zeitfilters.

vorausliegender Laute auf das folgende zu untersuchende Segment, um die Ermittlung der zeitlichen Dauer von Lautübergängen und Lauten für verschiedene Zwecke technischer und psycholinguistischer Art, so z. B. auch für die Herstellung des Ausgangsmaterials zur Durchführung von Lückentesten, bei denen die semantische Wirkung des inhaltlichen oder auch lautlichen Kontextes auf den Grad seiner Mächtigkeit, eliminierte Textelemente zu ergänzen, untersucht wird. Unser Mitarbeiter J. Kyritz hat dafür folgende Variante des Segmentierens entwickelt:

Fig. 1 zeigt das grobe Prinzip. Die 5 Funktionseinheiten bedeuten: 1 = Signalpeicher (Tonbandschleifengerät), wobei die Speicherzeit T_n von der Länge der Bandschleife und der gewählten Bandtransportgeschwindigkeit abhängt. 2 = Startimpulsgeber, der den Beginn jedes neuen Bandschleifenumschlages, also den Zeitpunkt t_0 durch Abgabe eines Impulses signalisiert. 3 = Impulsverzögerung, ein Zeitglied mit der Aufgabe, einen Ausgangsimpuls bei einem frei wählbaren Zeitpunkt t_m abzugeben. 4 = Schalter, der den Verstärkerweg vom Bandschleifengerät zum Abhörlautsprecher, gesteuert durch die vom Startimpulsgeber und vom Impulsverzögerer einlaufenden Impulse, ein- und ausschaltet bzw. umgekehrt. 5 = Verstärker mit Lautsprecher oder Kopfhörer, der durch den Schalter auf- und zugetastet wird.

Die Untersuchung eines speziellen phonetischen Problems fordert von dem zu verwendenden Zeitfilter höchstmögliche Genauigkeit in bezug auf Festlegung und Rekonstruierbarkeit der Schaltzeitpunkte, die außerdem zeitlich verschiebbar und meßtechnisch eindeutig erfaßbar sein sollen. Diese Forderungen lassen sich durch die Anwendung digital gesteuerter Impulsverzögerer und durch synchrone Speicher

nung der steuernden Impulse zusammen mit dem zu untersuchenden Signal verwirklichen.

Fig. 2 zeigt den detaillierten Aufbau schematisch an.

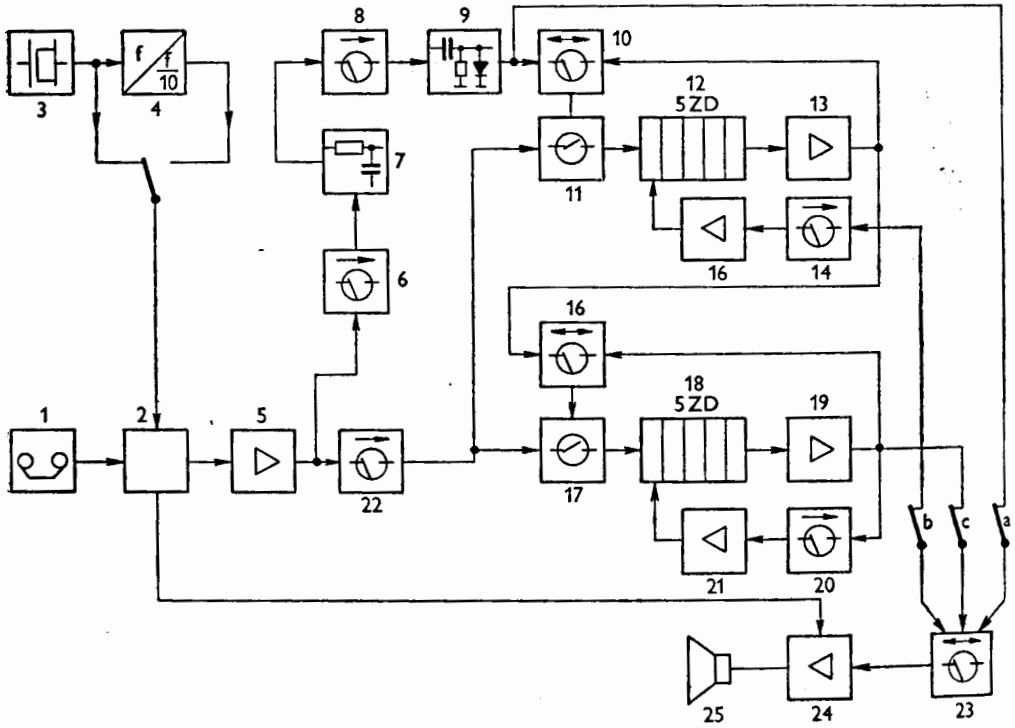


Fig. 2. Funktionsschaltbild eines im Institut für Phonetik und Kommunikationswissenschaft der HU Berlin entwickelten digital gesteuerten Zeitfilters mit optimaler Genauigkeit.

Legende:

1 Tonbandgerät, von dem die zu untersuchende Signalfolge auf das Bandschleifengerät überspielt wird. 2 Bandschleifengerät. 3 Temperaturstabilisierter quarzgesteuerter 10 kHz-Frequenzgenerator. 4 Frequenzuntersetzer 1 : 10. 5 Steuersignalverstärker. 6 Impulsformer (Monovibrator). 7 Integrationsglied. 8 = 6. 9 Differenzierglied. 10 Steuerbivibrator. 11 Torstufe. 12 5 hintereinandergeschaltete Zähldekaden. 13 Ausgangsimpulsverstärker. 14 Impulsformer. 15 Impulsleistungsverstärker. 16 = 10. 17 = 11. 18 = 12. 19 = 13. 20 = 14. 21 = 15. 22 = 14. 23 Schaltbivibrator. 24 Getasteter NF-Verstärker. 25 Lautsprecher.

Bem. Da der Bauelementaufwand bei herkömmlich aufgebauten Zähldekaden relativ hoch ist, wurden zwei Zeitfiltervarianten mit Schältröhren entwickelt. Variante 1 benutzt pro Dekade ein Dekatron mit direkt gekoppelter Ziffernanzeigeröhre (Z 572 S und Z 560 M, Hersteller: VEB Werk für Fernseh elektronik, Berlin, DDR). Variante 2 verwendet gasgefüllte dekadische Kaltkathodenzählröhren mit schräggestellten Kathoden (Typ N 4, Hersteller: Deutsche Glimmlampengesellschaft Pressler, Leipzig).

NASALE UND ORALE VOKALE — STRUKTUR UND PERZEPTION

GOTTFRIED MEINHOLD*

Nach ihrer spektralen Struktur sind die nasalen Vokale gegenüber den oralen vor allem durch eine Schwächung der Formanten F_1 , F_2 und F_3 (Smith, zit. bei Fant), durch das Auftreten von zusätzlichen Intensitätskonzentrationen oberhalb 1 kHz, durch einen Antiformanten bei 500 Hz (Hattori, Yamamoto und Fujimura) sowie durch eine relative Zunahme der Grundtonamplitude charakterisiert. Wenig beachtet wurde in der Literatur bisher der geringere Betrag der Gesamtintensität; Nasalvokale zeigen verminderte spezifische Schallfülle. Die genannten spektralen Erscheinungen — nämlich Abbau vokaldifferenzierender Teiltongruppen, Auftreten neuer oder vermehrter gemeinsamer Intensitäten vor allem im Bereich zwischen den Formanten — bewirken eine Nivellierung der Formantstruktur, eine Abtragung des scharf umrissenen Formantprofils. Der Grad der Konzentration der Intensität auf bestimmte Frequenzbereiche nimmt ab, die Intensitätsdispersion erhöht sich. Im Vergleich zu den entsprechenden oralen Vokalen büßen die nasalierten dadurch erheblich an spezifischer Klangcharakteristik ein und damit an Fähigkeit, untereinander gut wirksame Kontraste zu bilden; sie verlieren an Gestalt, weil der Grad ihrer Strukturiertheit abnimmt, ihre Identifizierung wird erschwert.

Normale Vokale ($[a, e, i, o, u]$) und ihnen entsprechende, im Sinne einer Rhinolalia aperta nasalierte, und zwar mit einem Grundton von ca. 113 Hz (männl. Vp), versuchte Vf. zunächst in ihrer Intensitätsdispersion mit Hilfe der Shannonschen Entropie zu vergleichen. Die Negentropie dient unter bestimmten Voraussetzungen als allgemeines statistisches Streuungsmaß. Demzufolge berechnet man das spektrometrische Bild mit den Frequenzklassen auf der Abszisse und der Intensität auf der Ordinate als relative Häufigkeitsverteilung und erhält für den Kompaktheitsgrad (Halle und Jakobson) eine Maßzahl, die als Vergleichswert für Vokale dienen kann, welche mit gleichen Spektrometern und gleichen Stimmen bei gleichmäßiger Aussteuerung und gleicher Grundfrequenz gewonnen wurden. Die spektrale Entropie dient also nicht als absoluter klanganalytischer Parameter, sondern als Vergleichsmaß in relativ engen Grenzen. Die Zunahme der Intensitätsstreuung bei nasalen Vokalen kann man gut an den Quotienten ablesen, die aus beiden Werten zu bilden sind

* Abt. Sprechwissenschaft der Universität, Jena (DDR).

$\left(\frac{H_{nas}}{H_{or}}\right)$. Der Vokal *a* wurde durch die nasopharyngeale Öffnung akustisch nur gering beeinflusst, zeigt also eine geringe Zunahme der spektralen Entropie, bei *i* und *o* dagegen ist sie erheblich.

Tabelle 1

	<i>a</i>	<i>e</i>	<i>i</i>	<i>o</i>	<i>u</i>
H_{oral}	4,2138	3,8923	3,4123	3,4923	4,0675
H_{nasal}	4,2389	4,3249	4,0356	4,1598	4,5681
$\frac{H_{nas}}{H_{or}}$	1,01	1,11	1,18	1,19	1,12

Der Entropiequotient blieb für *e*, *i*, *o*, *u* auch bei Vokalspektren von Hattori, Yamamoto und Fujimura, die Vf. miteinander verglich, in den Grenzen $>1,09$ und $<1,19$. — Die Berechnung der spektralen Entropie könnte u. a. als Objektivierungsmethode für Behandlungserfolge bei Rhinolalia bzw. Palatolalie (-phonie) dienen.

Im Perzeptionsversuch wurden zunächst Segmente derselben Vokale zu je 0,05 bis 0,06 sec Dauer von phonetisch vorgebildeten Hörern dekodiert (somit standen den Hörern über die Vokale ihrer Muttersprache hinaus Urteilsklassen zur Verfügung). Die Kürze sollte die Hörunsicherheit steigern. Die Registrierung der Imitationsprodukte bzw. der Urteile erfolgte durch den Vl.

Von oralen Vokalen wurden [*a e i*] sicher gehört, [*o*] dagegen teils als ,*u*' mit der geringen Urteilsunsicherheit von $H = 0,4999$ bt, [*u*] andererseits teils als ,*o*' und ,*m*' ($H = 0,5920$ bt).

Bei den nasalierten Vokalen bleibt [*ã*] sicher. In der Konfusionsmatrix (Tab. 2) werden daher nur die Urteile von [*i ã õ ù*] aufgeführt. Die aus den relativen Häufigkeiten der Konfusionsmatrix berechneten Urteilsentropien lauten (bei einer maximal zehn Urteilsklassen entsprechenden Entropie von $H_{max} = 3,3220$ bt):

	<i>i</i>	<i>õ</i>	<i>ù</i>	<i>ẽ</i>
H	2,3627	1,9152	1,8021	1,4464

(in bit)

Allerdings sagen diese Zahlen nichts über die qualitative Verschiebung bei der Dekodierung aus. [*i*] wird nämlich dann mit dem geringsten Verfehlungsindex (Tarnóczy) gehört, der durch die Summe der relativen Häufigkeiten aller anderslautenden Urteilsklassen ausgedrückt wird. (Da [*i*] zu 32,2 % i-ähnlich gehört wird, beträgt der Verfehlungsindex $100 - 32,2 = 67,8$ %.) Die aus dem Verfehlungsindex sich ergebende Rangfolge lautet also *i*, *ù*, *õ*, *ẽ* (vgl. Tab. 2).

Tabelle 2 Konfusionsmatrix (Angaben in %)

Urteile	, <i>e</i> '	, <i>i</i> '	, <i>o</i> '	, <i>u</i> '	, <i>y</i> '	, <i>æ</i> '	, <i>ø</i> '	, <i>m</i> '	, <i>a</i> '	, <i>é</i> '
<i>ẽ</i>	6,7					29,4	58,9	5,0		
<i>ĩ</i>	13,0	32,2		25,2			7,8	15,7	6,1	
<i>õ</i>			13,3	15,6		12,2	34,4		11,1	13,3
<i>ũ</i>				17,2	6,9	5,8	56,3	13,8		

Eine weitere Möglichkeit, den Kontrastverlust der abgehörten Nasalvokale untereinander zu messen, ergibt sich, wenn man die Vokale paarweise miteinander konfrontiert und die relativen Häufigkeiten der Urteile, die das jeweilige Vokalpaar gemeinsam hat, summiert (vgl. dazu auch Tab. 2, der die Zahlen entnommen sind):

Tabelle 3

Stim.	Gemeinsame Urteile	In %	Stim.	Gemeinsame Urteile	In %
<i>ẽ</i>	<i>e/a/m</i>	70,6	<i>ĩ</i>	<i>a/u/a</i>	39,1
<i>ĩ</i>	<i>e/a/m</i>	36,5	<i>õ</i>	<i>a/u/a</i>	84,6
<i>ẽ</i>	<i>a/æ</i>	88,3	<i>ĩ</i>	<i>u/m/a</i>	48,7
<i>õ</i>	<i>a/æ</i>	46,6	<i>ũ</i>	<i>u/m/a</i>	87,3
<i>ẽ</i>	<i>a/m/æ</i>	93,3	<i>õ</i>	<i>a/u</i>	51,0
<i>ũ</i>	<i>a/m/æ</i>	75,9	<i>ũ</i>	<i>a/u</i>	73,5

Aus den Mittelwerten der drei für jeden Nasalvokal ermittelten Prozentzahlen ergibt sich folgende Rangfolge der Kontrastaufhebung: [*i*] 41,4 %; [*õ*] 60,7 %; [*ũ*] 78,9 %; [*ẽ*] 84,1 %.

Derselbe Versuch, mit Vokalsegmenten doppelter Dauer (0,12 sec) wiederholt, ergab eine nur gering verbesserte Identifizierungssicherheit. Die Versuche werden mit längeren Segmenten fortgesetzt. Der eklatante Kontrastabbau der Nasalvokale bei Kürzung könnte als perzeptives Argument neben anderen (lautphysiologischen) Argumenten zur Erklärung von Tendenzen zur Denasalisierung in rezenten Sprachen (Hindi, Frz., Poln.) dienen. Hinsichtlich der kommunikativen Leistungsfähigkeit — soweit sie auf der Kontrastfähigkeit der Lautsignale beruht — kann man eine Rangfolge von den optimalen geschlossenen Vokalen über die sog. offenen bis zu den perzeptiv ungünstigsten nasalierten Vokalen aufstellen. Man muß also die perzeptiven Grenzen der Phonemunterscheidung ergänzend zu *Winkels* Untersuchung (1964) in engster Abhängigkeit von der spektralen Struktur der jeweiligen Vokale bzw. Laute definieren.

- Fant, G.: Acoustic Theory of Speech Production. s'Gravenhage 1960.
Halle, M. u. Jakobson R.: Grundlagen der Sprache. Berlin 1960.
Hattori, Sh., Yamamoto K. und Fujimura, O.: Nasalization of Vowels and Nasals. *Bull. Kobayasi Inst. Physic. Res.* 6 (1956) 226—235.
Tarnóczy, A.: Verständlichkeits-Stabilität, Konfusionsmatrix und Verfehlungsindex. *ZPSK* 17 (1964) 333—338.
Winkel, F.: Perzeptive Grenzen der Phonemunterscheidung. *Proceed. 5th Intern. Congr. Phonet. Sc.*, Münster 1964. Basel/New York 1965, 582—588.

DISCUSSION

Mangold:

1. Warum wurden die Vokale [I Y U Ō] nicht mitverglichen?
2. Vom Gesichtspunkt der Sprachuniversalien ist Entnasalierung nicht häufiger als Nasalierung.

Meinhold:

Warum nicht andere Vokale außer *a, e, i, o, u* untersucht wurden, hat vor allem arbeitstechnische Gründe. Es galt, zunächst an Hand einiger Kardinalvokale durch möglichst viele Entropieberechnungen die Verwendbarkeit des Entropiemaßes prinzipiell zu testen. Umlaute tauchen lediglich als Urteilklassen auf. Sicherlich könnte die erschwerte Unterscheidbarkeit der Nasalvokale auch ein Grund für ihre geringe Häufigkeit in den Sprachen der Welt sein, auch wenn hierfür ebenfalls lautphysiologische Gründe geltend gemacht werden können. — Die Versuche sollen bis zur Auffindung neuer spezifischer perzeptiver Grenzen der Phonemunterscheidung fortgesetzt werden, also neben geschlossenen und nasalen Vokalen auch mit sog. offenen.

A SCALING TECHNIQUE IN AN EXPERIMENT WITH VOWEL-LIKE SOUNDS

G. L. MEINSMA*

In the investigation I am going to describe to you, a group of 30 listeners were asked to determine, whether they thought that the second of a group of three artificial vowels was more like the first or more like the third vowel presented to them. 140 groups of three artificial vowels were presented to the listeners, the first and the third artificial vowel being fixed and the second being variable. The fixed points originated from the vowel triangle, as it can be drawn for the Dutch vowels (see fig.1),

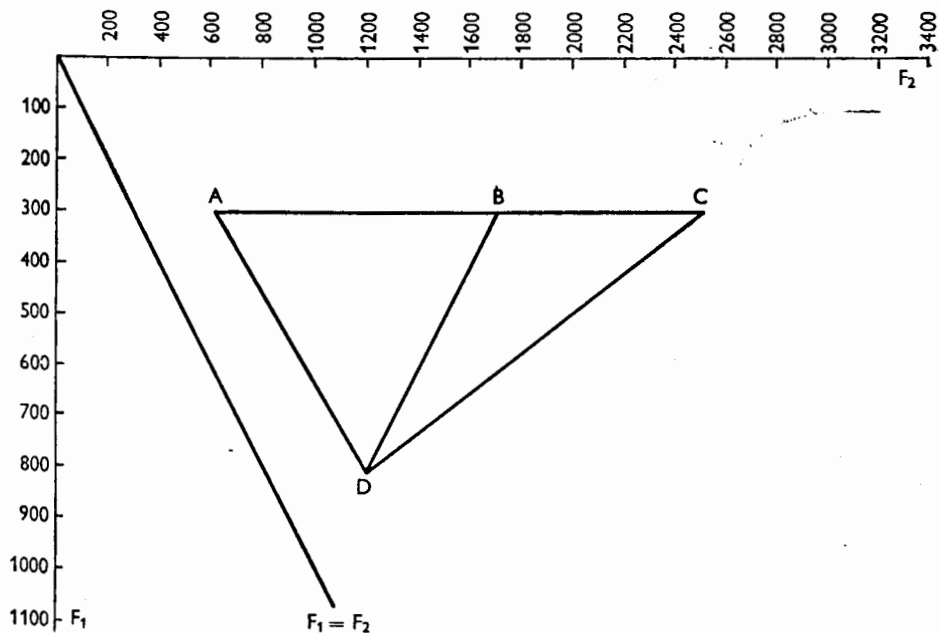


Fig. 1.

thus furnishing us with four scales: AD, BD, CD and AC. 33 points equally divided along each of these scales were taken as the variable vowel-like sound. The stimuli were presented in such a way that the first variable in a scale had the fixed vowel-like

* University of Amsterdam, Institute of Phonetic Sciences.

sounds—for example in the scale AD—in a sequence AD, the second next variable in a sequence DA, the third variable had again AD and so on. Every variable had a reversed position of the surrounding fixed points as compared with the pair preceding and the pair following. This applies to any of the scales mentioned. The pairs of three vowel-like sounds coming from the four different scales were presented in random order. The listeners were asked to score their opinion in a linear scale (see fig. 2). The first vowel-like sound—a fixed point in the formant scale—has its posi-



Fig. 2.

tion at the extreme left of the scale, the third vowel-like sound, being also a fixed point in the formant scale, at the extreme right of the same scale. The subjects were instructed that also the position within one of the seven parts of the scale was of importance. The listeners were not told that they were going to hear artificial vowels. The instruction mentioned only three sounds. The subjects were recruited from a department in which no information about vowels, vowel-systems and the vowel triangle was given.

When listening to an unfamiliar vowel one is inclined to relate this vowel to a known vowel class. We wanted to gain some insight into the grounds on which subjective judgments as to vowel difference or vowel resemblance are made. Therefore we took as our starting point the problem, to what degree vowel-like sounds, the formants of which are quite near one another, might be judged to be different. The backbone of this problem is the question whether a distance, that could be expressed in formant frequencies, could be scaled and related to these frequency distances. In other words: are physical distances correlated with perceptual distances?

In order to produce the vowel-like sounds mentioned above, we used a vowel generator, consisting in a pulse generator and two LCR-chains. The damped oscillations produced were summated and controlled as to damping coefficient, the amplitudes of the two formants produced and as to the respective frequencies of F1 and F2. The pulse generator, simulating the pulses given by the vocal cords, was adjusted at a frequency of 160 c/s. Pulse shape, damping coefficient and amplitude were set in such a way as to bear optimal resemblance to these parameters as they occur in actual Dutch vowels. The artificial vowels were recorded on tape at a same level and at electronically controlled distances in time. Every group of three vowel-like sounds was recorded twice at the same fixed distance in time and separated from the preceding and following groups by another pause of longer duration, which was also electronically controlled.

The subjects were isolated in boxes. The subjects were given printed, carefully worded instructions, allowing control by the experimenter.

Nevertheless some of them did not succeed in responding in the correct way. The

responses of the subjects were made on preprinted forms, allowing a quick coding for processing on an electronic computer.

As a first step we tested the hypothesis that scaling is possible and that the subject's responses rise monotonously with the stimuli.

We therefore applied Kendall's rank correlation test.

Schaal 15

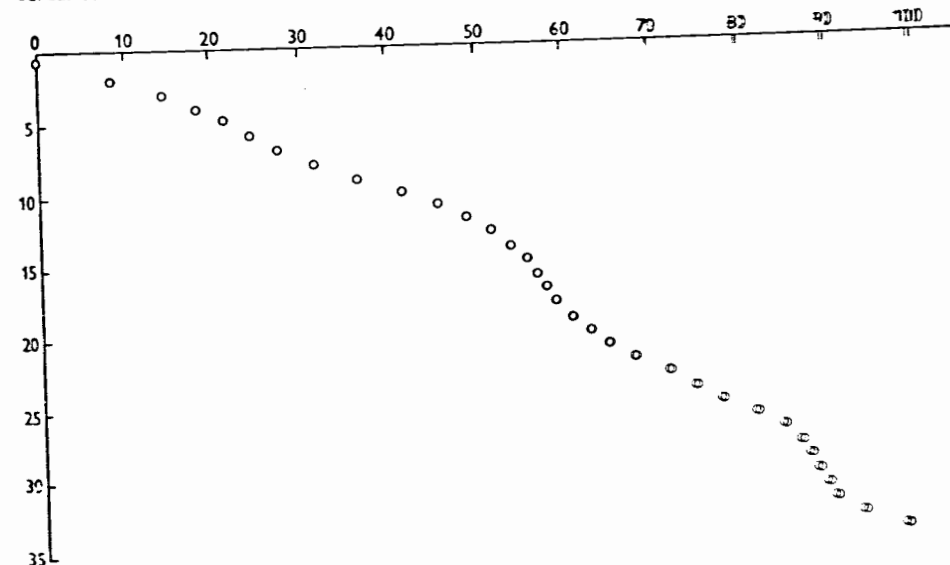


Fig. 3. Smoothed curve scale 1 (AD).

Schaal 25

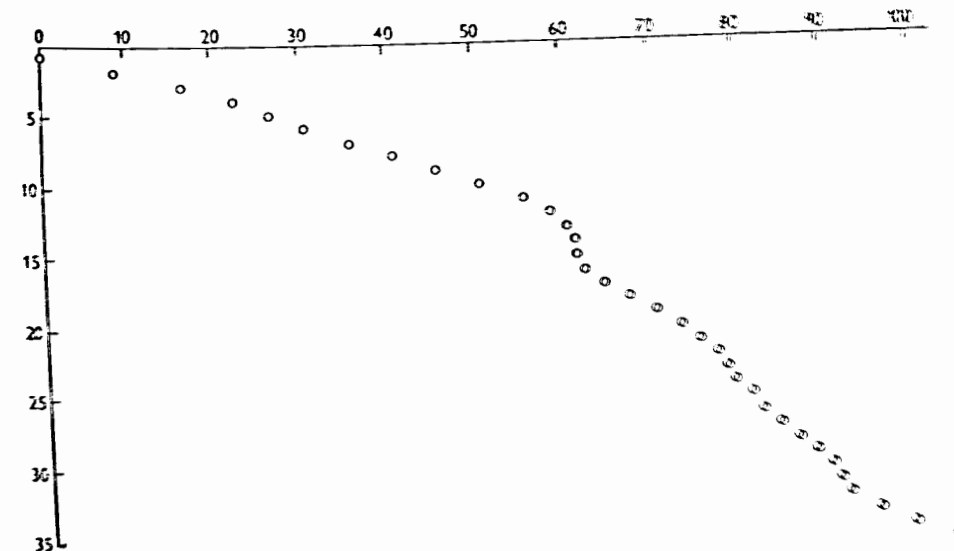


Fig. 4. Smoothed curve scale 2 (BD).

Out of 30 subjects 25 produced rank correlation coefficients sufficiently high to conclude to positive ranking within a 99.5 percent reliability.

5 subjects had very low or even slightly negative rank correlation coefficients. So that in their case we could not conclude significant ranking. The responses of these 5 subjects were therefore discarded.

Schaal 35

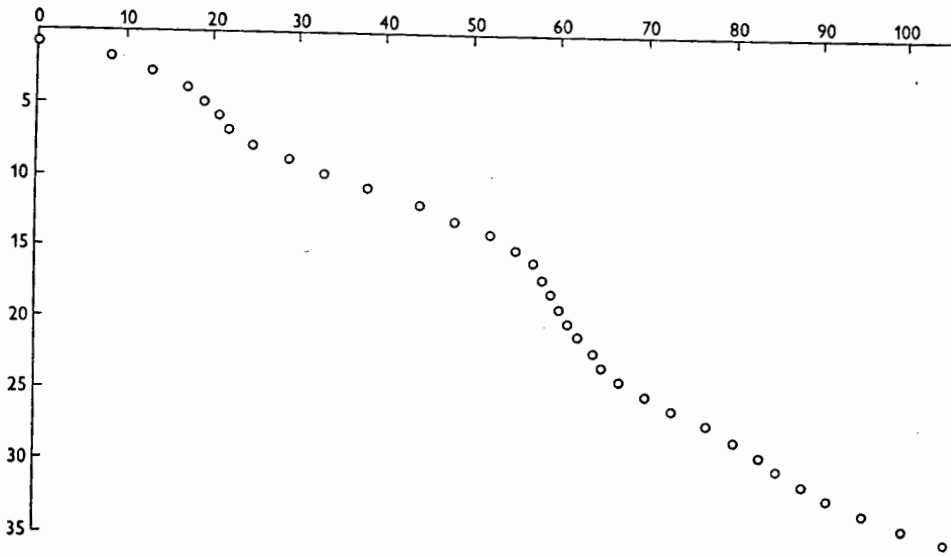


Fig. 5. Smoothed curve scale 3 (CD).

Schaal 45

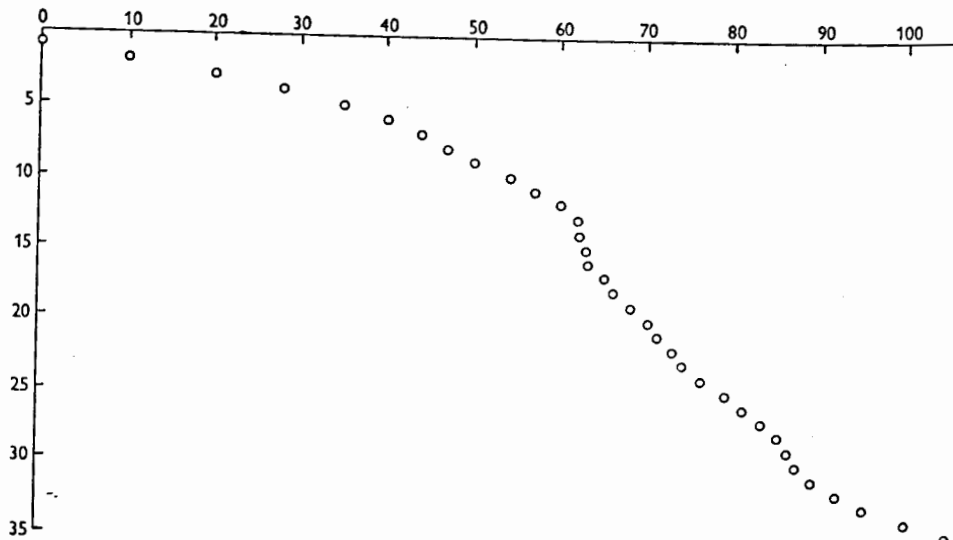


Fig. 6. Smoothed curve scale 4 (AC).

The accuracy of the scaling is expressed by the fact that the standard deviation for the response positions is about one seventh of the length of the whole scale. This applies to all scales and all stimuli.

Although the sequences in the perceptual and the physical scale were strikingly correlated, there was no tendency towards a linear relation. In our experiment—just as in so many other scaling experiments—our subjects showed a reluctance to score in the extremes of the scales.

Furthermore our subjects showed a tendency to score high in relation to a linear scale. Thus in the scale AD subjects responded more in the direction of D than is justified by the position of those stimuli in the physical scale.

On the raw data a process of digital filtering was performed in order to obtain smooth curves. (Fig. 3, 4, 5, 6.)

It is justified to speak about a perceptive vowel triangle.

If we map the physical vowel triangle on the perceptive triangle we notice that some areas are preferred, while others are avoided. The distribution of these areas seems to be related to the distribution of the Dutch vowels in the perceptual triangle.

Whether a native vowel system plays a role in the evaluation of perceptual distance between vowel-like sounds can only be established by repeating our experiment with subjects with different mother tongues.

Further investigation in this field is in progress.

DISCUSSION

Scully:

Please clarify the relationship between your results from two-formant stimuli and a chart of the first and second formant frequency in Dutch' which does not include all the parameters needed to specify vowel qualities acoustically.

Meinsma:

ad Scully: Principally this experiment was designed to investigate whether there is a relationship between the perception of two-formant stimuli and a chart of the first and second formant frequency in (Dutch) vowels.

Whether two or more formants are needed to specify vowel qualities acoustically was not the question. We may however refer to *Phonetica*, 15, 1966, p. 65—85 (J. G. Blom and J. Z. Uys; Some notes on the existence of a 'universal concept' of vowels) from which article it appears that in Dutch vowel-contrasts are sufficiently well described in two formants.

W. MERLINGEN

In dem alten Streit zwischen sogenannten Genetikern und Akustikern haben die älteren Phonetiker (Jespersen u. a.) mehr und stärkere Argumente zugunsten der artikulatorischen Seite gefunden. In neuerer Zeit neigt man aber immer mehr dazu, der *akustischen* Seite den Vorzug zu geben. Das wird oft betont, und zwar gerade auch in der Phonologie. So stellt z. B. Šaumjan (Problemy strukturnoj lingvistiki, 1962, 7) als eine Grundthese den Satz auf: „Phoneme sind akustische Elemente“; ähnlich sagt Pilch (Phonemtheorie I, 1964, 31f.): „Wir hören und verstehen sprachliche Rede . . . mit Hilfe unserer Ohren. Nur was wir *hören*, kann deshalb sprachlich von Belang sein“. Wenn sich auch schon Gegenstimmen erhoben haben, so wird doch dieser Aspekt noch weiter begünstigt durch die hohe technische Entwicklung akustischer Meßapparaturen, über die sich die Phonetiker so freuen. Angesichts dieser Entwicklung erscheint es angezeigt, die Aufmerksamkeit auf die Frage zurückzulenken, was denn hier eigentlich untersucht wird und worum es sich bei den Sprachlauten vom *sprachlichen* Standpunkt aus handelt. Wird mit den Apparaten wirklich das gemessen, was für die Sprache relevant ist?

Wenn wir fragen, was wir bei den Sprachlauten wahrnehmen, so scheint die Antwort zunächst klar zu sein: Geräusche, getragen von Schallwellen, so wie wir bei der Musik Klänge wahrnehmen usw. Aber in Wirklichkeit ist beim Sprachlaut mehr im Spiel. Zwei fundamentale Unterschiede sind zunächst greifbar:

1. Wir *hören* die Sprachlaute nicht nur, wir *machen* sie auch *selbst*; aktiv und passiv gehen wechselseitig ineinander über.

2. Das Erlebnis der Sprachlaute — ob gehört oder gesprochen — ist mit einem zweiten Erlebnis eng verbunden: mit dem Erlebnis der anderen Seite der Sprache, der Sprachinhalte, Vorstellungen, Denkelemente oder wie man sie nennen will.

Das sind Vorgänge, die ihren Schauplatz an ganz anderer Stelle haben als der Vogelschrei, die Musik oder alle sonstigen Geräusche, die wir hören. Wir haben ein eigenes *Organ* für die *Sprachlaute*, das nicht mit dem für die Geräusche schlechthin identisch ist, dieses Organ „kann“ sozusagen *mehr* als das Organ für die gewöhnlichen Geräusche, und es liegt *anderswo*; was das Sprachlautorgan von dem einfachen Geräuschorgan trennt, liegt tief *im Unterbewußtsein*, ist unserer direkten Beobachtung entzogen — was darin vorgeht, kann nur *erschlossen* werden. Dies ist es jedoch

gerade, worauf es beim Sprachlaut ankommt, die physikalischen Vorgänge ringsum geben uns keine Auskunft über die sprachlichen Vorgänge, nicht darüber, was im Sprachlaut-Organ vor sich geht. Wir müssen hier allerdings unsere Frage etwas genauer fassen. Sie darf nicht bloß lauten „was nehmen wir bei den Sprachlauten wahr?“, sondern „was nehmen wir mit unserem Sprachlaut-Organ wahr?“, oder, um die Unbewußtheit deutlicher herauszustellen: „was nimmt unser Sprachlaut-Organ wahr?“

Die Antwort ergibt sich fast von selbst schon aus der Überlegung, daß im Sprachlaut-Organ eine genaue Entsprechung zwischen Gehörtem und Gesprochenem bestehen muß; was man hört, muß mit dem, was man artikuliert, völlig identisch sein und umgekehrt — sonst könnte ja das Ganze nicht funktionieren. Es ergibt sich also die zunächst einigermaßen befremdliche Tatsache — befremdlich für die Eindrücke, die wir von unseren Sinnesorganen her gewöhnt sind —, daß wir nämlich mit unserem Sprachlaut-Organ die *Artikulationen* wahrnehmen, das Sprachlaut-Organ *hört* sozusagen die *Bewegungen der Sprechwerkzeuge*. Daß man Bewegungen hört — und nicht sieht —, das befremdet uns, wie gesagt, aber wir kommen um diese Tatsache nicht herum, die Sprache könnte auf andere Weise eben nicht funktionieren. Auf diese Weise erlernt ja auch das Kind die Artikulationen seiner Muttersprache: es hört Artikulationen und es reproduziert dann eben diese Artikulationen. Um am frühesten und besten funktionieren gerade die verstecktesten Artikulationen, wie die des Kehlkopfs, die dem Phonetiker und dem Sprachwissenschaftler immer noch Rätsel aufgeben, nicht aber dem nachsprechenden Kind. Auch als Erwachsener kann man ja den gewaltigen Unterschied zwischen dem direkten Anhören und dem theoretischen Kennenlernen eines Sprachlautes an sich selbst beobachten, bei groben wie bei feinen Schwierigkeiten. Direkt wirksam ist hier die Wahrnehmung durch das (unterbewußte) Sprachlaut-Organ; unser Wissen und Wollen ist dagegen meist so gut wie nutzlos.

Zu den Fragen der *Sprachwissenschaft* wird also die akustische Phonetik direkt nichts beitragen können, höchstens indirekt und ganz am Rande, etwa zur Kontrolle bei schwer zugänglichen Artikulationen. Völlig neu zu erarbeiten wäre darnach z. B. die Theorie des Phonems. Der Kern der Natur der Sprachlaute ist m. E. nur bei den Artikulationen, bei den Bewegungen und Stellungen der Sprechwerkzeuge zu treffen. Dies ist ja auch in der Kausalkette der innerste Punkt; alles andere — Schallwellen usw. — ist sekundär und tertiär. Von entscheidender Bedeutung ist dies bei der diachronischen Betrachtung, also bei den Fragen nach den sog. *Lautveränderungen*. Nie werden wir etwa, um ein augenfälliges Beispiel herauszugreifen, die eigenartigen und z. T. recht gegensätzlichen Akzentverhältnisse im Baltoslavischen verstehen und erklären können, solange wir dabei von *Tonhöhen* reden; erst wenn wir fragen und erforschen, um welche *Bewegungen der Kehlkopfteile* es sich handelt, und welche Bewegungen mit welchen anderen Bewegungen etwa verwechselt werden können und konnten, werden wir in diese Geheimnisse eindringen können. Es kommen hier Vertauschungen von hoch und tief vor, aber doch wohl nicht auf Grund von *Ver-*

wechslung von hoch und tief (was akustisch eben unmöglich ist); sondern durch *Verwechslung* von (ähnlichen) *Betätigungen*, die die umgekehrten akustischen Wirkungen haben. — Man sollte sich nicht durch äußerliche Folgeerscheinungen von der Natur der Dinge ablenken lassen, sondern ihnen möglichst *nahe* zu kommen trachten.

РЕЛЕВАНТНЫЕ ПРИЗНАКИ ПРОСОДИИ СЛОГА В СОВРЕМЕННОМ ЛИТОВСКОМ ЛИТЕРАТУРНОМ ЯЗЫКЕ

Е. МИКАЛАУСКАЙТЕ—Б. СВЕЧЕВИЧЮС—А. ПАКЕРИС*

Проблема просодии слога в современном литовском литературном языке очень актуальна. Данный вопрос не является еще решенным как в фонетическом, так и в фонологическом плане.

Хотя литовские слоговые акценты (литовск. *priegaidės*, англ. *Accents*) уже давно обратили на себя внимание лингвистов, побудив некоторых из них к поискам признаков речи, в которых закодирована различительная способность литовских акцентов, но и до сих пор не существует детального описания физико-акустических параметров этого явления.

В классической литуанистике (а также и в общем языкознании) многими поддерживается мнение, что акценты долгих слогов литовского языка, основанные на изменении частоты основного тона (F_0), имеют словоразличительное значение. И все же более детальное обследование диалектов литовского языка показало, что данное фонологическое свойство не распространяется на все говоры. Поэтому в настоящее время не все говорящие на литературном языке и не во всех долгих слогах пользуются фонологической оппозицией акцентов по указанному фонетическому признаку.

В литовской акцентологии указывается на два типа долгих ударных слогов: слоги с акцентом 1 (нем. *gestoßener Ton*, литовск. *tvirtapradė priegaidė*-знак ['] и слоги с акцентом 2 (нем. *geschliffener Ton*, литовск. *tvirtagale* — знак [~])**. Полагалось, что все слоги с акцентом 1 противопоставляются слогам с акцентом 2 по одним и тем же физическим признакам, т. е. движению основного тона (F_0), интенсивности (I), длительности (t) (в отдельности или комплексно). Тем не менее в исследованиях разных ученых индикаторами акцента считались не одни и те же акустические

* Государственный педагогический институт, Вильнюс.

** Некоторые исследователи (Р. Экблом, Г. Геруллис, П. Скарджюс, В. Вайткявичюте, В. Гринавецкис) в кратких слогах находят и акцент 3 (литовск. *trumpinė priegaide*), который, по нашему мнению, нельзя считать фонологически самостоятельным, так как долгие слоги противопоставляются краткими по ингерентным, а не по просодическим признакам речи.

признаки слога. Различия взглядов на этот вопрос можно выразить так:

- а) F_0 (Ф. Куршат, Ф. Фортунатов, Н. Трубецкой),
- б) $F_0 + I$ (А. Баранаускас, К. Яупюс, К. Буга),
- в) $F_0 + I + t$ (Р. Экблом, Г. Геруллис, П. Скарджюс),
- г) $I + t$ (А. Лайгонайте).

Наши эксперименты показали, что слоговая структура является главным индикатором, который указывает на фонологическую релевантность или иррелевантность акцентных признаков в литовском языке.

Детальные акустические исследования слогов монофтонгической структуры нас убедили в том, что различающими признаками являются движение частоты основного тона, интенсивности и относительная длительность, но, как показали результаты аудирования, вышеуказанные различающие признаки акцентов не являются признаками различительными (В. А. Артемов), так как аудиторы очень редко (лишь фонетически тренированные или выросшие среди населения западных диалектов) различают (вне контекста) слова типа *sú:ris* „сир“ и *sū:ris* „соленность“. Это свидетельствует о том, что в литературном языке акценты данного типа теряют фонологическую значимость. К классу этих слогов необходимо отнести и те слоги, которые содержат в себе дифтонги (*ie*) и (*uo*).

Акценты 1 и 2 в слогах дифтонгической структуры (*au*), (*ai*), (*ei*), характеризуются не только различиями в движении основного тона, интенсивности и длительности компонентов дифтонгов, но и очень своеобразными спектральными характеристиками. Следует отметить, что спектральным признакам свойственна инвариативность, в то время, как частота основного тона, интенсивность и длительность проявляют очень большую вариативность, зависящую от фразовой интонации, темпа речи и т. п. Первые компоненты дифтонгов (*av*), (*ai*), (*ei*), в слогах с акцентом 1 противопоставляются первым компонентам этих же дифтонгов в слогах с акцентом 2 по признакам напряжений-ненапряжений, закрытый-открытый, компактный-некомпактный (или диффузный). Определенные спектральные различия отмечаются и во вторых компонентах этих дифтонгов. Но, как показало аудирование, решающую роль при дифференциации акцентов 1 и 2 имеет спектральная структура первых дифтонгических компонентов. Следовательно, акценты в слогах данного типа противопоставляются по ингерентным признакам дифтонгов. По сравнению со слогами монофтонгической структуры, в этих слогах (т. е. в слогах, включающих в себя дифтонги) акцент 1 фонологически противопоставляется акценту 2 значительно более строго (при различении слов типа *áušta* „остывает“ и *aũšta* „светает“ получено 96, 96,06 % правильных ответов аудиторов).

Можно предположить, что фонологическое противопоставление акцентов

1 и 2 в слогах семидифтонгической структуры (*a, e + l, m, n, r*) также базируется на ингерентных признаках.

Наконец, следует выделить и третью акцентную слоговую структуру — слоги с дифтонгами (*ui*), (*oi*), (*eu*) и с семидифтонгами (*i, u + l, m, n, r*), которая занимает промежуточное положение (как по фонетическим, так и по фонологическим признакам) между слогами с долгими монофтонгами, дифтонгами (*ie*), (*uo*) и дифтонгами (*au*), (*ai*), (*ei*), дифтонгическими сочетаниями (*a, e + l, m, n, r*).

ANALYSE QUANTIQUE DES CAS DIFFICILES

R. F. MIKUS*

Comme vous avez déposé au vestiaire, avant d'entrer dans cette salle, vos manteaux et parapluies, vous êtes priés d'y laisser aussi les notions de voyelle et de consonne. Ce sont là des fantômes inventés par l'antiquité; ils vous seront tout à fait inutiles dans les analyses que je vous propose.

Ce qui existe en matière de phonétique linguistique, c'est l'expiration et les divers traitements auxquels elle est soumise par le jeu des organes de phonation. Par les mouvements des muscles intercostaux, l'expiration est d'abord portionnée en unités que j'appelle *quanta phonétiques*. Chaque quantum est porteur tout d'abord d'une *qualité* ou *couleur caractéristique* dénommée par l'antiquité „voyelle“; cette qualité peut avoir ou non des *marges* ou *couleurs marginales* appelées par la tradition „consonnes“. Les „voyelles“ et les „consonnes“ ne sont pas des entités „coexistant“ avec l'expiration, „l'accompagnant“ ou „véhiculées“ par elle, ce ne sont que les noms donnés par l'antiquité à certaines propriétés de l'expiration. Etant un segment infinitésimal de l'expiration, le quantum est l'unité physico-physiologique du langage; il a, outre les „couleurs“, aussi d'autres propriétés: quantité ou longueur intensité et hauteur musicale relatives. Je ne vous fatiguerai pas en essayant de vous convaincre que cette explication ne porte pas seulement sur la terminologie, mais qu'elle concerne le fond même de la phonétique.

Ma communication est intitulée: Analyse quantique des cas difficiles. Par ce terme j'entends les cas dont la phonétique traditionnelle donne une explication insuffisante. Mon temps étant limité, je me borne à exposer l'exemple de *Sprung*.

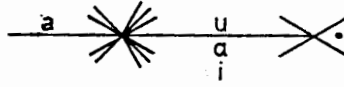
Sprung est un quantum à couleur caractéristique *u* introduite par le chaînon explosif *špr-* et couverte par la marge implosivo-explosive finale *-ng'* ou *-nk'* (dessin). La figuration traditionnelle CCCVCC est trompeuse aussi parce qu'elle fait croire que les couleurs caractéristiques et les marges sont, du point de vue de la durée, égales. Il n'en est rien: les marges sont réalisées par l'implosion et l'explosion et sont par conséquent momentanées, tandis que les couleurs caractéristiques comportent une durée rationnelle, car elle sont réalisées par les tenues. Au lieu de parler de „voyelles“ et de „consonnes“, il convient mieux de parler des phases phonatoires

* U.O.C., Lubumbashi.

(implosion, tenue, explosion) *sans égard à la qualité* (couleur) réalisée. On se demande où est la „voyelle“ lorsque je change dans ce dessin la couleur caractéristique en bleu (*Sprang*) ou en rouge (*Spring*).

La chaîne parlée est constituée de quanta accommodés l'un à l'autre. *Ein Sprung* comporte deux quanta dont le premier, coloré en *a*, est clos par le chaînon implosif *yn* et rattaché directement au chaînon explosif *špr-* (dessin).

ein Sprung



a yn š p r u n k'

>><<< ><.

A MODEL OF THE VOWEL SYSTEM OF DUTCH

H. MOL*

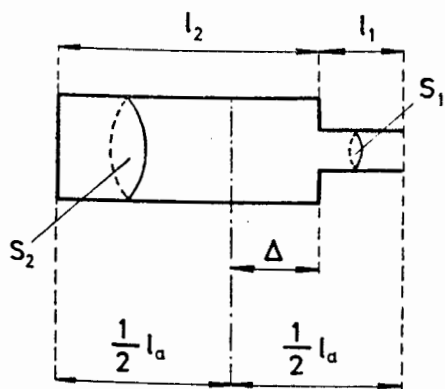
At the spring-meeting of the Acoustical Society of America at Washington in 1965 I discussed the backgrounds of the twin-tube model of the vocal tract. I pointed out how to transform the well-known formant-formula

$$\tan \frac{\omega l_1}{c} \tan \frac{\omega l_2}{c} = \frac{S_1}{S_2} = k$$

into

$$\cos \frac{\omega l_a}{c} = \frac{1-k}{1+k} \cos 2\Delta \frac{\omega}{c}$$

by introducing a new coordinate, the *eccentricity* Δ that describes the position of the joint as its distance to the middle of the twin-tube.



$$\operatorname{tg} \frac{\omega l_1}{c} \operatorname{tg} \frac{\omega l_2}{c} = \frac{S_1}{S_2} = k$$

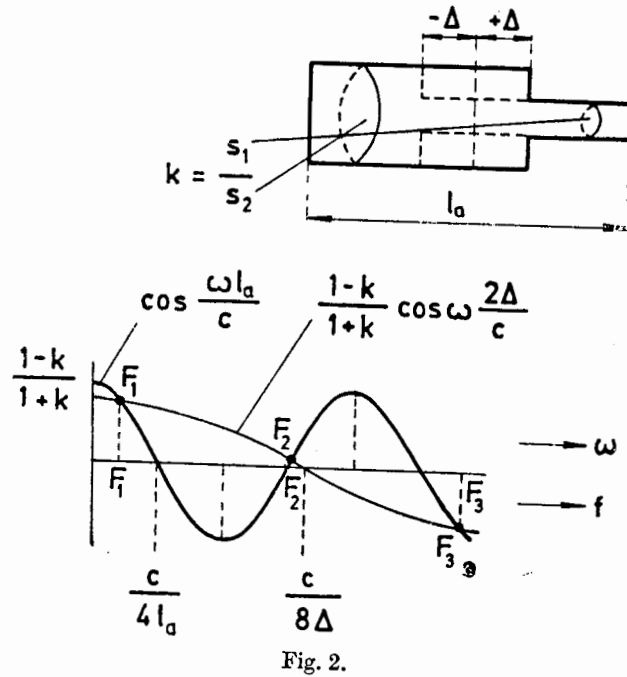
$$\cos \frac{\omega l_a}{c} = \frac{1-k}{1+k} \cos \omega \frac{2\Delta}{c}$$

Fig. 1.

The cosine formula can be solved either graphically or by means of a computer. In describing the vowels of Dutch we used both methods: the simple graphical method in order to get a quick insight into the nature of the vowel diagram and the computer method for more sophisticated applications.

* Institute of Phonetic Sciences of the University of Amsterdam.

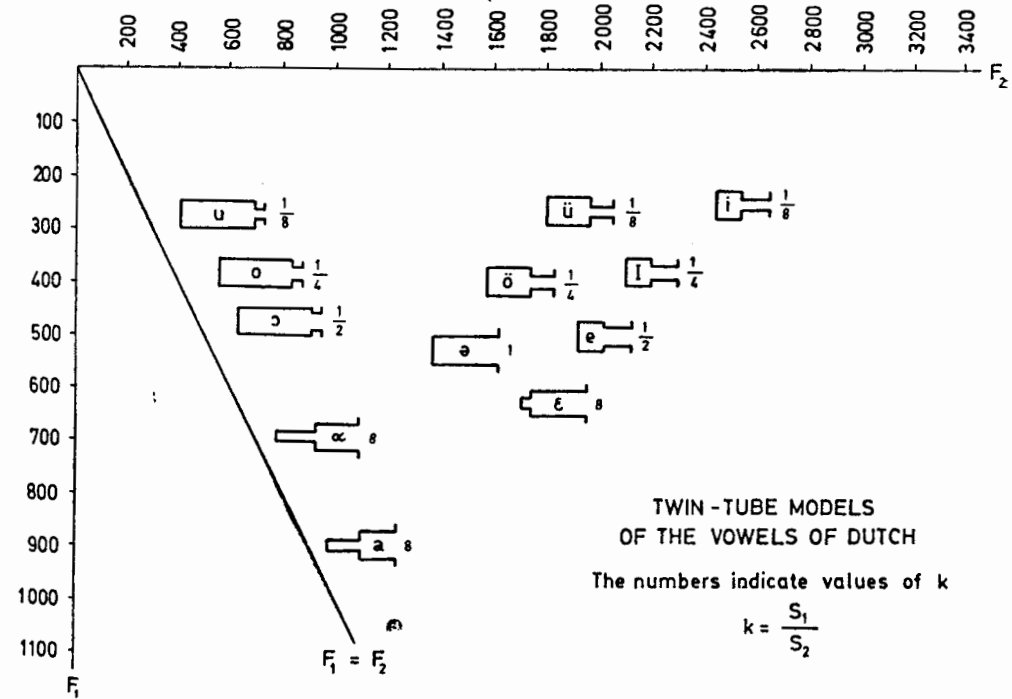
The geometry of a twin-tube is completely determined by its parameters l_a , Δ and k . The 12 vowels of Dutch in isolated key words, carefully pronounced by a native speaker, can be described by a set of 12 twin-tubes producing the same formants. The parameters of each twin-tube can be chosen on basis of the study of X-ray photo-



graphs, palatograms and formant-measurements of that speaker. By now it is well-known that the absolute formant positions — hence the absolute values of the parameters—do not and indeed cannot function phonemically. Instead, it is the contrasts between the formants of the vowels that form a system. To put things more precisely, though the *absolute* formant positions may vary between speakers, the *relative* positions of the vowels in the complete vowel diagram in the $F_1 F_2$ plane are always the same for native speakers of the same language, provided these speakers have no apparent speech defects and are free from outspoken dialectical peculiarities. The formant positions depicted in the slide refer to a talker with about average anatomic dimensions. We may pin the phonemic system on him as well as on any other normal speaker of Dutch.

The acoustical length of the vocal tract of a speaker depends on the vowel he pronounces in an isolated word. For instance, in [i] the vocal tract is short, because, acoustically, it begins somewhere behind the teeth, the larynx is usually somewhat lifted, whereas the back of the tongue is low. On the other hand, in [u] the lips are protruded, the larynx is in a low position, whereas the back of the tongue is high. In that case the axial acoustical length of the vocal tract is at its maximum. In [a]

the length is about average. It suffices to confine ourselves to 3 different contrasting lengths of the vocal tract in Dutch for depicting the phonemic system, but this is not a decisive factor as long as the contrasts are being conserved. As to the parameter



$k = \frac{S_1}{S_2}$, five suitably chosen values will do. For the eccentricity Δ , three contrasting values suffice.

Strictly speaking, the twin-tube model is only a mathematical concept involving several idealisations. It is very tempting, and also very rewarding, to *materialize* the twin-tube model in the shape of two hard-plastic tubes, driven by a suitable sound source imitating the real larynx. For this artificial larynx we have chosen an acoustical siren, visible on slide 4. Compressed air is being let in through a triangular opening in series with a rotating slit. In that way a continuous series of air puffs, having the necessary shape, is generated. Each air puff begins gradually but ends with an abrupt cut-off which generates powerful damped oscillations in the artificial vocal tract. The thus-generated vowels do not sound any sillier than the sustained vowels produced by life speakers.

In Dutch there are sequences of two vowels that are traditionally called diphthongs by many, though they are nothing but vowel clusters in which both elements carry phonemic information. We shall not deal with these diphthongs here.

Moreover, there are vowel clusters in which only the first element carries phonemic information, whereas the second element is merely an audible transition without a phonemic function. This type is often called a diphthongal vowel. We shall not discuss it here.

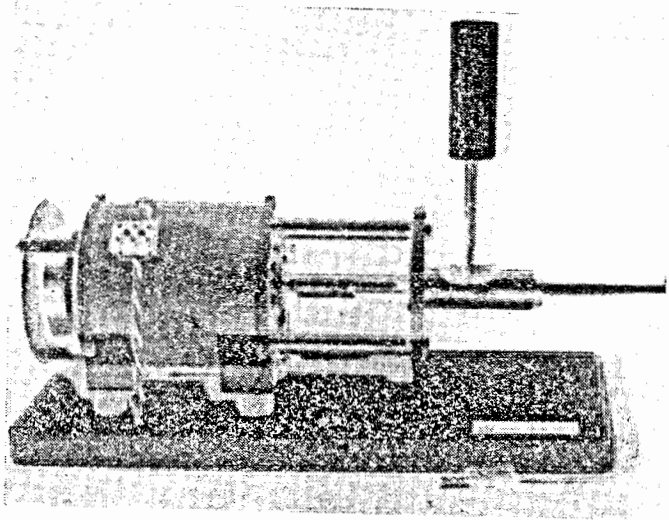


Fig. 4.

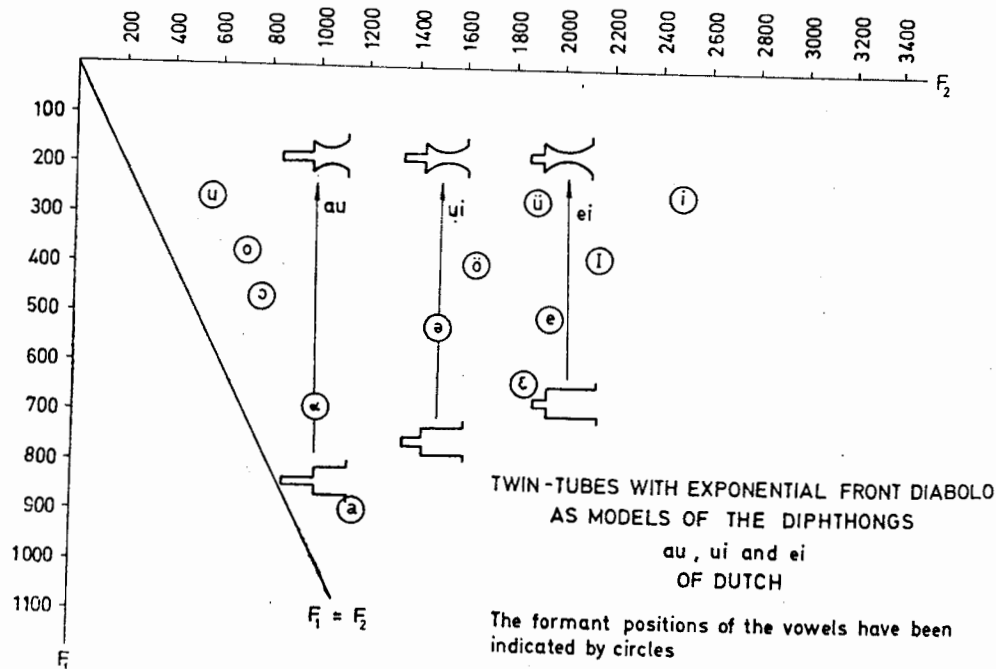


Fig. 5.

In Dutch, however, there is a special class of diphthongs of a *dynamic* nature, namely the speech sounds [au], [ui] and [ei]. They consist of two vowel-like elements having the same (or slightly different) F_2 , gliding into each other by a rapid shift of F_1 .

It is very probable that this shift, together with the position of F_2 , carries the phonemic cue. Especially towards its end, however, this shift is so rapid that it seems to fall out of step with the comparatively slow articulatory movements. For the description of these three diphthongs we want a model that produces a practically constant F_2 and an ever-accelerating glide of F_1 . Such a dynamic model can be derived from the twin-tube model by squeezing its mouth-tube into the shape of a diabolo. For reasons of mathematical simplicity we suppose that the diabolo consists of two identical exponential horns. In that case squeezing can be reduced to simply changing the flare of the horn.

The time allotted to this paper does not allow me to present the convincing calculations but they will be published elsewhere.

СОПОСТАВИТЕЛЬНЫЙ АНАЛИЗ АКУСТИЧЕСКИХ СВОЙСТВ АНГЛИЙСКИХ И ЛАТЫШСКИХ ДИФТОНГОВ

М. А. НЕЙЛАНД*

Согласно классификации дифтонгов Л. В. Щербы в изложении М. И. Матусевич (Введение в общую фонетику, Л., 1948, стр. 60) дифтонги английского (Ая) и латышского языков (Ля) принадлежат к двум различным типам дифтонгов. Дифтонги Ая принадлежат к ложным нисходящим дифтонгам, т. е. только один из его компонентов, а именно первый ударный образует вершину слога. Второй компонент — безударный, ненапряженный и вершину слога не образует. Дифтонги Ля являются образцом истинных дифтонгов, оба компонента которых слоговые и произносятся отчетливо и с одинаковой напряженностью.

Ая и Ля имеют хорошо развитую систему дифтонгов: в Ля — 10 дифтонгов — (ai), (au), (ei), (eu), (oi), (ou), (iu), (ui), (ie), и (uo), в Ая дифтонгов — (ei), (ou), (ai), (au), (ɔi), (ɔa), (ra), (ɛa), (ua).

Пять дифтонгов в этих языках имеют сходный фонетический состав: английские (er), (ou), (ar), (au), (ɔr) и латышские (ei), (ou), (ai), (au), и (oi).

Два английских дифтонга (ra) и (va), несмотря на их различный звуковой состав, произносятся подобно латышским дифтонгам (ie) и (uo).

Дифтонги Ая до сих пор исследовались главным образом в артикуляционном аспекте. Природа скользящей артикуляции английских дифтонгов и соотношения их компонентов зарубежными фонетистами трактуется противоречиво.

Исследование дифтонгов Ля также можно считать завершенным только в артикуляционном аспекте. Поэтому главной задачей данного экспериментального исследования является изучение акустических свойств дифтонгов Ая и Ля и их сопоставление.

В процессе исследования были использованы следующие методы: а) анализ языкового материала; б) электроакустический анализ основных физических характеристик; в) метод спектрального анализа и г) метод кино съемки артикуляции губ.

* Латвийский государственный университет, Рига.

Согласно данному исследованию, английские дифтонги следует разделить на две группы:

1. устойчивые дифтонги, напр., (er) и (ou) и
2. скользящие дифтонги, напр., (ar), (au), (ɔr), (ɛə), (ɪə) и (uə).

В устойчивых дифтонгах оба компонента артикулируются полностью. В начале и в конце артикуляции устойчивых дифтонгов происходят более медленные изменения в артикуляции, чем в середине. В скользящих дифтонгах качество нервного компонента устойчиво, но второй компонент является глайдом. Поэтому в конце артикуляции этих дифтонгов изменения в артикуляции происходят быстрее.

Все латышские дифтонги являются устойчивыми. Средняя долгота (er) — 317 мс, а (ou) — 322 мс.

В дифтонге (er) первый компонент в 1,1—2 раза длиннее второго компонента или оба компонента (er) одинаковой длительности. В дифтонге (ou) первый компонент в 1,3—1,7 раза длиннее второго компонента или второй компонент (ou) в 1,3—1,6 раза длиннее первого компонента. Длительность переходной части между компонентами дифтонгов (er) и (ov) одинакова и составляет в зависимости от позиции 10—25 % длительности данных дифтонгов.

В отличие от дифтонгов (er) и (ou) первый компонент дифтонгов (ar), (au) и (ɔr) значительно длиннее второго компонента: в дифтонгах (ar) и (ɔr) в 2 раза длиннее, в (au) — в 2—3 раза длиннее второго компонента. Длительность второго компонента дифтонга (au) обычно меньше длительности соответствующего компонента дифтонга (ar). Поэтому переходная часть между компонентами дифтонга (av) длиннее, чем в дифтонге (ar) и составляет 16—29 %, а в (ar) — 15—26 %.

В дифтонгах (ɪə) и (ɛə) соотношения длительности компонентов разные: первый компонент длиннее второго и наоборот, или оба компонента одинаковой длительности. В дифтонге (uə) первый компонент обычно длиннее второго компонента. Переходная часть между обоими компонентами данных дифтонгов выражена нечетко и составляет в зависимости от позиции 12—35 % длительности центральных дифтонгов.

В латышских дифтонгах, которые артикулируются с тремя видами слоговой интонации: длительной, нисходящей и прерывистой первый компонент обычно немного длиннее второго.

Сопоставительный анализ акустических свойств дифтонгов Ая и сходных по фонемному составу дифтонгов Ля с нисходящей слоговой интонацией позволяет заключить, что в обоих языках нет ни одной пары дифтонгов с равным комплексом физических свойств, т. е. направлением частоты основного тона, интенсивности, а также качественными и количественными соотношениями компонентов.

Это подчеркивает различный характер артикуляционных особенностей дифтонгов обоих языков.

Английские дифтонги (er), (ar), (au) и (ɔr) являются восходяще-нисходящими или нисходяще-восходящими по направлению основного тона, но нисходящими по направлению интенсивности. Латышские дифтонги (ɛi), (ɔi), (ɔv) и (ɔi), наоборот, являются нисходящими по направлению основного тона, восходяще-нисходящими по направлению интенсивности.

Исключение составляют английские (ɪə), (uə) и латышские (ie) и (uo), в которых направление основного тона и интенсивности имеет две тенденции.

Интервал понижения частоты основного тона в дифтонгах Ля больше, чем в соответствующих дифтонгах Ая.

В дифтонгах обоих языков широкие компоненты больше отличаются друг от друга, чем узкие. Дифтонги Ая в среднем длиннее дифтонгов Ля.

Дифтонги Ая артикулируются протяжнее, ровнее (с меньшим интервалом понижения основного тона) и с небольшой модуляцией тона, а дифтонги Ля с более выраженным интервалом понижения основного тона и без его модуляции. В дифтонгах Ля второй компонент длиннее и яснее, чем соответствующий компонент в сходных дифтонгах Ая.

Однако в обоих языках имеются дифтонги, в которых некоторые физические свойства являются идентичными, например, направление основного тона, интенсивности, качество широкого или узкого компонентов, соотношения долготы компонентов и др. В этом случае артикуляция дифтонгов обоих языков воспринимается идентичной, но в процессе обучения недопустима идентификация сходных дифтонгов обоих языков.

Сопоставительный анализ обоих языков подтверждает их принадлежность к двум различным типам дифтонгов: английских — к ложным, а латышских — к истинным. В английском языке исключением является лишь дифтонг (ov), который артикулируется как истинный дифтонг, а в латышском языке — (ie) и (uo), которые произносятся и как ложные восходящие дифтонги.

По акустически артикулярным свойствам в обоих языках имеются устойчивые и неустойчивые, позиционно независимые и зависимые, лабиализованные и нелабиализованные дифтонги.

Данное исследование дает научно-теоретическую базу а разработке методики преподавания английских дифтонгов латышам.

DISCUSSION

Stojanović:

What are the grounds for assuming the Latvian two-segment sequences should be considered diphthongs, as both components were said to be syllabic?

How did you arrive at the conclusion that English (ei, ou) are the shortest diphthongs, and did you take into consideration the difference in length of the two diphthongs occurring before voiceless and voiced consonants respectively?

Neiland:

Latvian two-segment sequences are considered true diphthongs, as they are pronounced in one syllable, and have a single intensity culmination. They can be opposed to each other as one unit. As both components of Latvian diphthongs are distinctly and fully pronounced, linguists often refer to them as syllabic.

English diphthongs were experimentally investigated by us isolated and occurring between voiceless consonants. In these positions (ei) and (ou) are far shorter than other diphthongs.

SOME SUPRASEGMENTAL CRITERIA FOR CONTEXTUAL RELATIONS BETWEEN SENTENCES IN ENGLISH

GERHARD NICKEL

In the present state of sentence analysis it would seem a promising venture to reconsider some of the theses of the Prague School. Vilém Mathesius's dichotomy of "formal" and "functional sentence analysis"¹, for instance, takes on a new significance in the light of N. Chomsky's controversial notion of "underlying and surface structure".

Mathesius holds that to describe a sentence one must rigidly discriminate an underlying constituent structure, through which all relevant syntactic relations between the parts of the sentence are specified, and a kind of surface structure ("*Mitteilungsstruktur*"), through which certain aspects of the surface structure are determined. The latter is influenced both by the intention of the speaker and by the context. This notion seems to shed some light on the problem of word order, which has not been solved yet by the generativists.

In an interesting article K. E. Heidolph has shown that the arrangement of constituents in a German sentence can be largely derived from anaphoric relations with contextually adjoining sentences.² Other phenomena like selection of article, pronominalization, accentuation as well as certain aspects of 'complex sentence formation' can be explained in the same way.³ In this paper I shall attempt to describe some contextual conditions which must be regarded as relevant for an analysis of suprasegmental features. A formulation of exact rules cannot be attempted here owing to the absence of a complete grammar taking into account contextual regularities. Moreover, these rules would have to refer to the underlying structure of English sentences, a matter still much in dispute.⁴

¹ This terminology has been used by J. Firbas. Cf. the bibliography in J. Firbas's paper, 'Some Thoughts on the Function of Word Order in Old English and Modern English', *Sborník prací Filologické Fakulty Brněnské University* 6 (A5), 1957, pp. 72—100.

² K. E. Heidolph, 'Kontextbeziehungen zwischen Sätzen in einer generativen Grammatik', *Kybernetika*, 3 (1966), pp. 97—109.

³ Cf. G. Nickel, 'Some Contextual Relations between Sentences in English', to appear in: *Actes du Xème Congrès International des Linguistes*, Bucharest, 1968.

⁴ The following studies, while remaining within the framework of generative grammar, contain proposals deviating more or less radically from Chomsky's own version of the theory: B. C. Hall, *Subject and Object in Modern English*, M. I. T. Doct. Diss., Cambridge, Mass 1965; Ch. J. Fillmore,

It can be shown that an NP in a given sentence always carries the primary accent when introduced for the first time. If mentioned for the second time it can only carry a secondary accent. Thus if a sentence S_i contains an NP and is followed by S_j , which has NP¹ anaphorically related to NP, NP in S_i receives the primary accent (/'), while NP¹ in S_j is given the secondary accent (/'). Furthermore, an NP mentioned for the first time, is normally preceded by the indefinite article and often appears in final position, while NP¹ is preceded by the definite article (often realized as *this*) and usually has initial position. Using Vilém Mathesius's terminology, we shall call the former *rheme* and the latter *theme*. Hence we will be able to distinguish between a rhematic and a thematic accent:

(1) Bill had an apple. (2) This apple had red cheeks.

If there is no such NP in the antecedent sentence the rhematic accent is automatically given to the verb. (If the verb is accompanied by particles or certain modifications then these parts carry the accent.):

(3) Bill ate the apple.

The proper name in (3) is equivalent to an NP with the definite article additionally marked with the distinctive feature [+ unique]. We may assume that every definite anaphoric NP, even if marked [+ unique], permits the contextual elimination of the head NP. In this case an anaphoric pronoun appears, which can be considered as the form of the definite anaphoric article in isolation. Cf.: (4) *He ate the apple*. (4) has only one accent, viz. the rhematic one. Thus one may assume that the thematic accent as a suprasegmental feature can be eliminated together with the head NP. On the other hand the rheme does not allow pronominalization. Neither (5) nor (6) are equivalent to (1): (5) *John had something*. (6) *Bill had it*. (6) appears acceptable only if e.g. the neutral gender is contrasted with another gender.

Sentences like (7) *Bill ate the apple*, and (8) *Bill ate the apple* can appear in the same context as (3). They do, however, imply additional contextual conditions. Cf.: (9) *Bill (not Jim) ate the apple*. (10) *Bill ate the apple (not the pear)*. It has been pointed out before that sentences like (7) and (8) presuppose others which are often identical except for the constituent carrying the *contrastive accent*.⁵ These sentences, as it were, correct antecedent sentences paradigmatically.⁶ They are particularly often used in

⁴ 'A Proposal Concerning English Prepositions', *MSLL*, 19, Washington, D. C. 1966, pp. 19—33; G. Lakoff, *On the Nature of Syntactic Irregularity*, Report No. NSF-16, Harvard University, Cambridge, Mass., 1966; J. Lyons, 'Towards a "Notional" Theory of the Parts of Speech', *JL*, 2 (1966), pp. 209—236; P. M. Postal, 'On So-called Pronouns in English', *MSLL*, 19, Washington, D. C. 1966.

⁵ M. Bierwisch, 'Regeln für die Intonation deutscher Sätze', *Studia Grammatica*, 7 (1966), pp. 151f.

⁶ Cf. K. E. Heidolph, *op. cit.*

dialogue.⁷ We may assume that the *wh*-question demands equivalent contextual conditions. It differs from the answer given to it only in the NP to which the *wh*-marker refers. This assumption would correspond to the fact that any lexical constituent which can be made the object of a question can have the contrastive accent:

- (11) Who polished his shoes? — Jôhn polished his shoes.
 (12) What did John polish? — John polished his shoes.
 (13) What did John do with his shoes? — John polished his shoes.

The assumption that only lexical items can carry the contrastive accent seems to be invalidated by the following sentence: (14) *John did see the girl*. However, the sentence makes sense only if complemented, e.g. in the following way: (15) a) *John did not see the girl*. b) *John did see the girl*. A change of speaker is also involved. Therefore this type is very common in dialogue. The contrasting of a) with b) can be interpreted as a contextual explication of emphasis, which can refer to the assertive morpheme of the verb.

Either kind of special accentuation, contrastive as well as emphatic, often superimposes the contextually conditioned regularities of thematization and rhematization. Consequently, Heidolph in his article has explained contrastive and emphatic statements as synonymous repetitions of antecedent sentences. Every primary accent on a second-mention NP seems indeed to be interpreted as a contrast.

As to intonation, the accents, whose position are contextually determined, mark the points where changes in pitch relevant to the interpretation take place. If it is true that there are only two relevant 'Tonhöhenlagen's then they are likely to be determined by the two prominent points within a sentence, viz. rheme and theme. The thematic accent corresponds to a pre-ictic rising 'Tonbruch': $\uparrow\bar{x}$, which according to Isačenko and Schädlich announces a falling intonation at the end of the sentence. Without this falling intonation a sentence remains a fragment. This observation is supported by the fact that an NP with thematization never appears in isolation. This kind of fragmentary sentence differs clearly from the acceptable type of grammatical ellipsis.⁹ If sentences with only one relevant 'Tonbruch' are possible,¹⁰ then they must have the post-ictic falling one: $\bar{x}\downarrow$, which often appears at the end of a sentence, but in any case constitutes the final cadence of a question. This falling intonation may well be regarded as a very important, criterion of the rheme. A falling 'Tonbruch' can be easily isolated in the case of an ellipsis. In an answer to a *wh*-ques-

⁷ R. Gunter, 'On the Placement of Accent in Dialogue', *JL*, 2 (1966), pp. 159—179.

⁸ A. V. Isačenko and H.—J. Schädlich, 'Untersuchungen über die deutsche Satzintonation', *Studia Grammatica*, 7 (1966), p. 58. (Though the authors have only investigated German this seems to apply to English too.)

⁹ R. Gunter, 'Elliptical Sentences in American English', *Lingua*, 12 (1963), pp. 137—150.

¹⁰ A. V. Isačenko and H.—J. Schädlich, *op. cit.*, pp. 60f.

tion the NP to which the *wh*-marker refers appears in isolation with the falling 'Tonbruch': (15) *Who ate an apple?*

The bôy (ate an apple).

Finally a short remark on selective (= restrictive) and descriptive (= modifying) clauses. In a descriptive relative clause there are usually two rhemes, one of which is the head constituent while the other lies within the descriptive modifying clause. If the modifying clause is not reduced it is preceded by an optional pause:

(16) I'll take the first *tráin* (//) which stops at the main station.

This observation might lead us to the assumption that every sentence can have only one rhematized constituent unless there is a sentence pause between the two clauses in question. Here we might have an additional criterion for defining more clearly the descriptive relative clause whose embedding is less complete than that of the restrictive clause. Very often a change in strategy is noticeable in spoken texts.¹¹ The integration of the following clause is carried through only half way: (17) *I'll take the first tráin. Which stops at the main station.* This is hardly possible with the selective type, which is closely integrated. Here the matrix NP does not carry a rhematic accent: it is closely connected with the selective clause, which carries the rhematic accent. Pauses in this case have to be interpreted in all likelihood as hesitation phenomena.

DISCUSSION

Wode:

Vielleicht könnte man den Begriff der Wiederholung in einigen Fällen durch die zusätzliche Unterscheidung von semantischen Unter- und Oberklassen etwas schärfer fassen. Z. B.

you'll find all sorts of cars: big vehicles, small vehicles, an old Ford, a new Chrysler, anything you like.

Die Oberklasse (*vehicle*) bleibt unbetont, die Unterklasse (*Ford* bzw. *Chrysler*) wird betont.

Firbas:

With regard to the problem of the relation between repetition and thematization raised by Dr. Wode in the discussion, I should like to emphasize the necessity of elaborating the criterion of what may be termed the *narrow scene*, i.e. in fact that of the very purpose of the communication. (I have touched upon this problem in *Non-Thematic Subjects in Contemporary English*, *Travaux Linguistiques de Prague* 2, Prague 1966, p. 246).

¹¹ B. Drubig, *Kontextuelle Beziehungen zwischen Sätzen im Englischen* (Kiel, 1967), M. A. Thesis (mimeographed). I feel deeply obliged to the author for many valuable ideas offered in his thesis and in long discussions. Helpful suggestions were also given to me by my Lektor and informant Roger L. Snook, M. A.

EINFLUSS DES LEE-EFFEKTES UND DES GERÄUSCHES AUF DIE SPRACHE

A. NOVÁK*—M. MORÁVEK**—M. PETŘÍK***—J. LANGOVÁ†

Die sprachliche Äußerung wird durch eine Reihe von Rückkoppelungsvorgängen kontrolliert. Dazu gehören die Propriozeption (Informationen aus den aktiven Muskeln), die akustische, durch die innere und äußere Luftleitung sowie durch die Knochenleitung vermittelte Rückkoppelung, die die Informationen über die akustischen Eigenschaften der Sprache überträgt, und schließlich die semantische Rückkoppelung, die von den oben erwähnten Mechanismen ausgeht und außerdem die Informationen verwertet, die durch die übrigen Sinne zugeleitet werden. Durch diese semantische Rückkoppelung wird der Inhalt der Äußerung und die Wirkung der sprachlichen Mitteilung auf die Umgebung kontrolliert.

Wenn diese Rückkoppelungsvorgänge in Tätigkeit treten sollen, müssen die entsprechenden Signale perzipiert werden. Eine Möglichkeit, wie man sich von der Art und Intensität der Perzeption überzeugen kann, ist eine künstlich hervorgerufene Störung im Prozeß der Perzeption.

In der Sprachforschung und in gewissem Maße auch in der Erforschung des Denkens benützt man in der letzten Zeit sehr oft den sogenannten Lee Effekt, dessen Prinzip darin besteht, daß der sprechenden Person mittels eines adaptierten Tonbandgerätes eine verspätete Information über ihre eigene Sprachproduktion ins Ohr übertragen wird. Diese Maßnahme beeinflußt die Sprache störend bei einem großen Prozentsatz normalsprechender Personen und verbessert im Gegenteil auffallend z. B. die Sprache der Stotterer.

Wir stellten uns die Frage, ob sich der Charakter der störenden Wirkung des Lee-Effekts verändert, wenn gleichzeitig ein Thermalgeräusch von verschiedener Intensität einwirkt. Soweit uns bekannt ist, wurden diese Untersuchungen bisher noch nicht durchgeführt.

Wir untersuchten insgesamt 20 gesunde Personen ohne Sprachstörung im Alter von 17—42 Jahren, 17 Männer und 3 Frauen in 8 verschiedenen sprachlichen Situationen: bei normalem Lesen, beim Lesen unter Einwirkung eines Geräusches von 60, 70

* ORL katedra IDVLF Praha (předn. doc. MUDr. K. Bláha).

** Ústav leteckého zdravotnictví, Praha.

*** Foniatriká laboratoř FVL Praha.

† Foniatriká klinika FVL Praha.

und 80 dB, weiter unter dem Einfluß des Lee-Effekts, wobei die Verspätung 0,145 sec betrug, und schließlich unter der gleichzeitigen Wirkung des Lee-Effekts und des Geräusches in der Intensität von 60, 70 und 80 dB. Die Versuchspersonen lasen eine Liste von 30 phonetisch ausgeglichenen Wörtern; davon waren 20 dreisilbig und 10 zweisilbig. Die angefertigten Tonbandaufnahmen übertrugen wir auf einen Oszillographen. Als Maß der Einwirkung der angeführten Einflüsse auf die Sprache bewerteten wir in der Aufnahme die Länge der einzelnen Wörter und die Pausen zwischen ihnen. Durch wiederholtes Abhören bewerteten wir weiter die Anzahl von Fehlern, die die Probanden machten.

Die Ergebnisse unserer Untersuchungen sind folgende:

1. Unter dem Einfluß des Lee-Effekts allein kommt ein statistisch hoch signifikanter Anstieg der Durchschnittslänge der Wörter sowie ein hoch signifikanter Anstieg der Gesamtzahl von Fehlern zustande.
2. Die gleichzeitige Wirkung des Lee-Effekts und des Geräusches in den oben angeführten Intensitäten verursacht eine praktisch gleiche Verlängerung der Wörter wie der Lee-Effekt allein. Im Gegenteil besteht kein statistisch signifikanter Unterschied zwischen dem normalen Lesen und dem Lesen unter Einwirkung des Thermalgeräusches allein.
3. Beim Lesen unter gleichzeitigem Einwirken des Lee-Effekts und des Geräusches in der Intensität von 80 dB kommt jedoch eine bedeutende Verminderung der Fehleranzahl zustande im Vergleich mit dem Lesen unter der Wirkung des Lee-Effekts allein.

Es besteht kein Unterschied zwischen der Wirkung der erwähnten Faktoren auf zwei- und dreisilbige Wörter.

Aus den angeführten Ergebnissen geht hervor, daß das Sprachsignal auf verschiedenen Ebenen des zentralen Nervensystems als komplizierte Struktur wahrgenommen wird. Bei der gleichzeitigen Einwirkung des Lee-Effekts und eines Geräusches von 80 dB, durch das das Sprachsignal um durchschnittlich 10 dB maskiert wird, sind die Zentralmechanismen imstande, den Klang der Stimme im Geräusch zu detektieren und mit Hilfe dieses Signals das Sprachtempo zu korrigieren, auch wenn die eigentliche Sprache schon ganz unverständlich ist. Die Länge der Wörter ist ähnlich wie bei dem Lee-Effekt allein vergrößert. Da jedoch der semantische Charakter der Signale durch das Geräusch völlig unterdrückt wird, wird das Artikulationsprogramm und seine Realisation durch die Zentralmechanismen nicht korrigiert und die Anzahl der Fehler kann praktisch vernachlässigt werden.

Das Geräusch allein in den Parametern, die wir untersucht haben, beeinflußt die Sprache nicht.

Subjektiv am schlimmsten wird die Situation bei gleichzeitiger Einwirkung des Lee-Effekts und des Geräusches empfunden, wenn beide in der Intensität von 70 dB einwirken. Wenn auch diese subjektiven Empfindungen nicht quantitativ ausgedrückt und zu keinen anderen Werten, die wir gemessen haben, hinzugefügt werden können, kann man diesen Umstand als sehr wichtig betrachten. In dieser Situation haben

Sprachsignal und Geräusch die gleiche Pegelgröße und deshalb kommt der Lee-Effekt und der störende, verzerrende Einfluß des Geräusches auf das Sprachsignal zur Geltung, welches das durch den Lee-Effekt „belästigte“ Zentrum viel schlechter verarbeitet.

Die Ergebnisse unserer Beobachtungen bestätigen unsere Hypothese, die wir im Jahre 1965 ausgesprochen haben, daß der sogenannte semantische Inhalt der Information, die bei der Einwirkung des Lee-Effekts durch die akustische Rückkoppelung übertragen wird, nicht den einzigen störenden Faktor darstellt. Die Verständlichkeit des durch den Lee-Effekt verspäteten Sprachsignals ist bei der Einwirkung des Geräusches von 80 dB 0,3 Prozent. Trotzdem wird die Länge des Lesens eines Probetextes ähnlich wie unter der Einwirkung des Lee-Effekts allein beeinflußt. Dies beweist uns, daß das Zentrum die Fähigkeit hat, aus einem Gemisch von 80 dB Geräusch und 70 dB Sprachsignal nur die Eigenschaften dieses Signals zu detektieren, die rückwirkend nur das Sprechtempo beeinflussen.

DÉFINITION DE LA SYLLABE ET DE SES COMPOSANTS (VOYELLE, DIPHTONGUE, SONORE, CONSONNES)

EUDOVÍT NOVÁK

Au lieu de la communication de *M. I. Stan, La nature de la syllabe*, j'ai averti les auditeurs des nouvelles définitions purement linguistiques de la syllabe et de ses composants, fondées par moi sur le canevas *commencement — centre — fin*, que voi-ci:

Syllabe — segment rythmique total le plus petit du corps phonique de la langue comportant commencement, centre et fin.

Voyelle — son susceptible, en signalant commencement, centre et fin à la fois, de former par lui-même une syllabe ou, en connexion avec consonne(s), susceptible de signaler à la fois soit commencement et centre, soit centre et fin, soit uniquement centre de la syllabe.

Diphthongue — union de deux voyelles ou d'une voyelle et d'une sonore, susceptible, dans certaines langues, en signalant commencement, centre et fin à la fois, de former par elle-même une syllabe, ou, en connexion avec consonne(s), susceptible de signaler à la fois soit commencement et centre, soit centre et fin, soit uniquement centre de la syllabe.

Sonore — consonne susceptible, dans certaines langues, en dehors de la fonction consonantique, de former soit seulement centre de la syllabe, soit en dehors de cela aussi, en connexion avec consonne(s), à la fois ou bien commencement et centre, ou bien centre et fin de la syllabe.

Consonne — son susceptible à lui-même ou en connexion avec d'autres consonnes de signaler uniquement soit commencement, soit fin de la syllabe.

(Pour l'argumentation visant ces définitions, ici améliorées, voir notre étude intitulée *Caractère périphérique des consonnes dans le système phonologique et dans la structure syllabique*, dans le II^e tome des *Travaux linguistiques de Prague*, paru en 1966).

THE ACOUSTIC FEATURES OF THE POSITIONAL VARIANTS OF NASALS IN CZECH

JIŘINA NOVOTNÁ*

In this contribution I want to concentrate on some positional variants of nasal consonants in Czech, from the point of view of their delimitation on the three dimensional spectrum.

When assessing the acoustic character of Czech nasals I base my observations on standard spoken Czech. The acoustic spectrum was made on two types of apparatus: 1. The sonagraph, i.e. a Visible Speech instrument adapted and produced by Kay-Electric Co., and 2. The Czechoslovak TESLA spectrometre which is a variant of the sonagraph.

Like all other sounds, Czech nasal consonants are modified in different environments and their basic acoustic features differ—more or less—from the basic type which occurs most often in the intervocalic position.

Czech nasals *m*, *n*, *ň* belong to occlusives, but differ from other occlusives by a weaker explosion and also by the visibility of F_3 . The strongest explosion among the occlusives is to be found with *ň* which is—accoustically—a variant of the palatal α and its explosion is in some position marked by noise characteristics. The nasals also differ from other consonants by the fact that their explosive character is manifested only in some positions, for instance in the intervocalic and initial positions. On the other hand there are many positional variants of the nasals which are often characterised on the spectrum by the absence of explosion. This concerns in particular the final position of the nasal *m* and *n* (while the palatal *ň* keeps the explosion in the final position with a number of speakers) or the positional variants of the nasal *n* before sibilants, affricates, labials, alveodentals and velars.

Let us observe more closely some of the individual variants. First of all I will discuss the nasal *n* before velars when the basic variant of the nasal changes in quality into η . This is a typical case of the assimilation of the place of articulation which has been described in almost all the phonetic works on the Czech language.

On the acoustic spectrum the difference between the two variants is usually manifested by the presence of explosion with the alveodental *n* and the absence of explo-

* Phonetic Laboratory of Institute of Czech Language, ČSAV, Praha.

sion with the velar η . With the velar η there is often a barely noticeable F_2 and always a visible F_3 . This acoustic difference between the alveodental n and the velar η is obviously the consequence of different articulation conditions. The alveodental n is shaped by one mouth closure: contact of the tongue on the palate arch mainly in the teeth ridge area. The explosion occurs after the occlusion is freed. With the velar η the articulation shifts from the alveodental area to the velum; in the usual pronunciation of the $n + k$ combination the off-glide is simplified or the off-glide of the velar η coincides with the on-glide of the velar k . The position of the organs of articulation with the occlusion of the two sounds is identical, with the only difference being in the position of the soft palate. After the closing of the passage into the nasal cavity, the occlusion is freed in the area of the soft palate.

The nasal n often has a similar picture in the combination $n + s + k$ which in the Czech language has a fairly high frequency of occurrence, mainly in the suffixes *-ský*, *-ská*, *-ské*.

The changed acoustical character of the nasal can be observed in these cases while listening to the progressing segments in the work with the segmentator when it is signalled by a strong nasalisation of the preceding vowel, an easy transition from the vowel to the consonant and a lack of explosion. The nasalisation of the preceding vowel is naturally caused by the fact that the soft palate descends and frees the entry into the nasal cavity already during the transition from the vowel to the nasal; the mouth occlusion either does not take place at all or is very weak. The pronunciation is of course potential, but the many examples found in my material testify to the high frequency of this variant of the nasal in a position preceding the combination $s + k$. We seem to be confronted here (as in the case of the combination $n + k$), with the influence of a velar, articulated in an entirely different articulation place from the basic type of the nasal. The acoustic picture of this variant of the nasal is close to the velar η and the pronunciation of none of my speakers showed a visible explosion. The nasalisation of the preceding vowel is often manifested by a clear F_3 vowel.

The qualitative change of non-palatal n into palatal \tilde{n} or at least of a palatalised variant was observed in the combination $n + \tilde{s} + \tilde{t}$. This seems to be the result of the softening influence of the palatal \tilde{t} . In this case too a weak nasalisation of the preceding vowel was observed, as well as the pronunciation of the nasal without explosion. The softening influence of the palatal \tilde{t} on the preceding nasal is of course strongest in the combination of $n + \tilde{t}$: this is the case of two consonants articulated in the same place of articulation. These groups are characterised by a strong softening of the nasal, manifested in the spectrum by the prolonged duration of the sound and a weakening of F_3 .

A weakened explosion of the nasal n can also be found in the position before the labials p , b . In sonagrams of this combination the explosion usually is not visible at all. Acoustically the variant of n in this position is manifested by weakened clarity of F_2 . On the other hand F_3 is clear almost in all cases. This picture of the nasal and

a short pause before the explosion of the labial as well as the weakened explosion all characterise this combination.

It should be added that the lack of explosion in the nasal in consonant clusters is manifested in a different way than with other occlusives. The acoustic spectrum of consonant combinations where one of the components is an occlusive shows very often that the basic acoustic feature of an occlusive, namely explosion, does not necessarily have to keep its priority and can be replaced by a secondary feature which then serves for the identification of the phoneme: for instance in the combination of two occlusives the explosion of the first component can be replaced on the spectrum by a pause before the explosion. With some variants of nasals, for instance before sibilants, even this secondary feature disappears.

НЕКОТОРЫЕ ВОПРОСЫ ИССЛЕДОВАНИЯ ИНТОНАЦИИ ТИПОВ ПРЕДЛОЖЕНИЙ В ТЮРКСКИХ ЯЗЫКАХ

А. Н. НУРМАХАНОВА*

Нами были исследованы некоторые вопросы, связанные с интонационными особенностями типов предложений в тюркских языках¹. В исследовании применялись экспериментальные методы, разработанные советскими учеными. В каждом из сравниваемых языков имеется ряд отличительных черт в образовании разнообразных типов предложений, сложившихся в результате неравномерного развития лексико-грамматических и интонационных средств, что обусловлено историческими факторами. В фонетическом отношении узбекский язык отличается от каракалпакского и казахского языков, утратой закона сингармонизма. Главными интонационными средствами выразительности в узбекском языке выступают увеличенное время произнесения и устойчивость частотного диапазона предложения, характерные для плавного изменения движения основного тона. Изменения в интенсивности произнесения для узбекского языка менее значимы, чем для каракалпакского и казахского языков. В казахском языке изменения движения основного тона выражены более ярко, время звучания слогов менее длительно, чем в каракалпакском языке. Например:

Ўндан х₁рамма к₁ўрқад₁ди дедингми? (узб.) — 1862 м/сек.

оннан х₁рамме к₁орқ₁ады дедиң бе? (карак.) — 1656 м/сек.

одан барлығы к₁орқ₁ады дедиң бе? — (казахск.) — 1425 м/сек.

Проведенное нами экспериментально-фонетическое исследование интонации повествовательных, вопросительных, побудительных и восклицательных предложений с узбекском, каракалпакском и казахском языках пока-

* ИЯЛ АН УзССР Ташкент.

¹) Нурмаханова А. Н. Типы простого предложения в тюркских языках (грамматические, интонационные особенности на материалах узбекского, каракалпакского, казахского языков), Ташкент, 1965.

Ее же, Типы предложений по модальности и интонации в тюркских языках. Рукопись докторской диссертации, Ташкент, 1966.

Ее же, Об экспериментальном изучении интонации в тюркских языках. „Общественные науки в Узбекистане“, 1964, № 11 и др.

зало, что посредством интонации предложение получает не только законченность, но и подразделяется на разнообразные типы.

Повествовательные предложения отличаются от других типов своими грамматическими и интонационными средствами, обеспечивающими сообщения в процессе общения. Повествовательным предложениям присуще восходяще-нисходящее движение основного тона на главноударном слоге.

Вопросительные предложения посредством лексических, грамматических и интонационных средств не только выражают вопросительную мысль говорящего, (стремление узнать о чем либо неизвестном), но и побуждают слушающего одновременно к ответу на вопрос. Интонация вопросительных предложений в тюркских языках приобретает весьма разнообразный характер в зависимости от места смыслового центра вопроса в предложении. Вопросительные предложения с вопросительной частицей имеют восходящее движение основного тона, при этом ядро вопроса обычно находится в самом конце предложения. Вопросительные предложения, образованные посредством вопросительных местоимений и наречий, находящихся в начале или середине предложения, имеют восходяще-нисходящее движение основного тона на главноударном слоге.

Побудительные предложения характеризуются восходяще-нисходящим движением основного тона на главноударном слоге. В отличие от повествовательных предложений, на заударных слогах наблюдается восходящее движение основного тона.

Исследование подтвердило самостоятельность восклицательных предложений, имеющих собственную грамматическую структуру и интонацию. Они являются собственно восклицательными, их нельзя превратить в другие типы предложений. Вместе с тем, экспериментальное исследование также показало возможность превращения других типов предложения (повествовательных, вопросительных, побудительных) в восклицательные. При этом необходимо учитывать ситуацию, при которой было произнесено данное предложение, также восклицательные предложения являются как бы несобственными, так как они обратимы.

Восклицательные предложения имеют восходящее движение основного тона на главноударном слоге. Они характеризуются ярко выраженной интенсивностью произнесения и расширенным диапазоном. Доказательством этому могут служить наши данные, полученные в результате сравнительного анализа узбекского, каракалпакского и казахского языков.

Таким образом, экспериментально-фонетическое исследование позволило вскрыть взаимосвязь речевой интонации в рассматриваемых языках со структурой предложения, с одной стороны, и гибкую, многогранную функцию интонации — с другой, выполняющую смысло-различительную роль, обусловленную ситуацией общения.

В связи с многогранной функцией интонации исследуемые типы предло-

жений отличаются друг от друга, служат не только средством сообщения мысли, но и выражают чувство, волеизъявление говорящего, побуждения и эмоции говорящего. При этом необходимо учитывать не только формально-грамматические и семантические, но и экстралингвистические факторы, выражаемые интонацией.

Исследование показывает, что сравнение двух или более языков дает широкие возможности находить те или иные лексико-грамматические и интонационные особенности каждого языка в отдельности и для всех систем тюркских языков, показывающих неравномерное их развитие, а также ряд общих закономерностей в образовании различных типов предложений.

DISCUSSION

Артемов:

Сравнительное изучение интонации казахского, узбекского и каракалпакского языков существенно не только для тюркологии, но и для общих проблем экспериментально-фонетических исследований, таких например как различительные признаки интонации.

SPEECH PRODUCTION AT THE NEURO-MUSCULAR LEVEL*

S. ÖHMAN*—A. PERSSON**—R. LEANDERSON***

The phonological output of a generative grammar may be regarded as an abstract description of certain subjectively observable aspects of speech production. The question as to what form the outputs of the phonological component should take is thus dependent on the nature of the physical parameters controlling speech production as well as on the properties of subjective phonetic observation. The purpose of the present paper is to emphasize the rather high degree of abstraction involved in even the lowest level phonological representations of human utterances

Modern acoustic phonetics has provided a firm understanding of the physical basis relating articulatory configurations to the resulting sound pressure wave. On the other hand, research at this level has demonstrated an overwhelming inconstancy in the acoustic correlates of the phonological invariants. Studies of articulatory dynamics suggest, however, that much of this variability may be due to built-in physiological properties of the articulatory organs and of the neural circuits controlling them. A model of speech production adequate for phonological purposes should thus incorporate a neuro-motor level exhibiting these properties.

An attempt to devise a partial model of this sort is summarized in Fig. 1. The upper left corner shows schematized sound spectrograms of the VCV utterances /ø:gø:/, /ø:ga:/, /a:gø:/, and /a:ga:/ as spoken by a male Swedish talker. These data have been extracted from a larger study described elsewhere.¹

Note that the formant transitions from the initial /ø/ into the medial /g/ are different when /ø/ and /a/ occupy the final position of the utterance. Similarly, the transitions from the medial /g/ into the final /a/ are different when /ø/ and /a/ occupy the initial position of the utterance. Thus, as was also observed by Menzerath and Lacerda,²

* Speech Transmission Laboratory, Dep. of Speech Communication, Royal Institute of Technology, Stockholm.

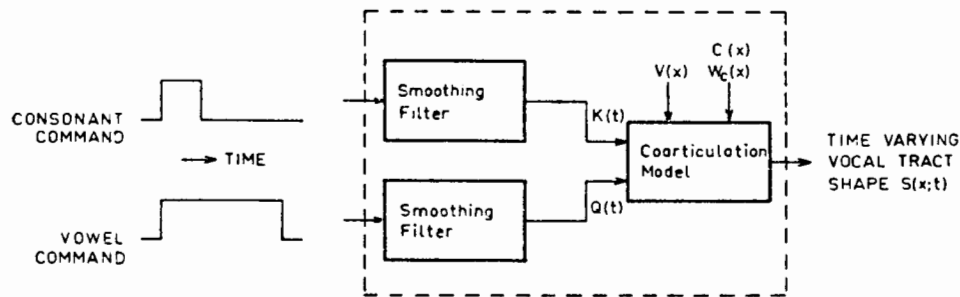
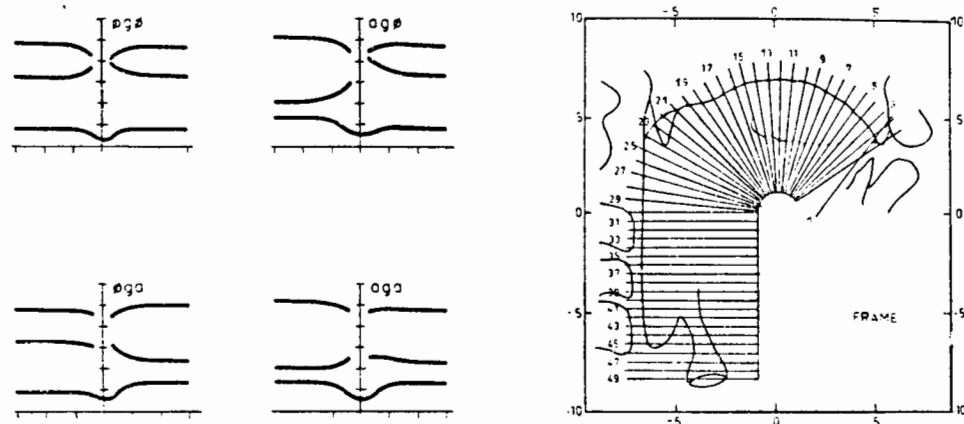
** Central Neurophysiological Laboratory, Karolinska Sjukhuset, Stockholm.

*** Phoniatrie Clinic, Karolinska Sjukhuset, Stockholm.

¹ Öhman, S. E. C.: "Coarticulation in VCV Utterances: Spectrographic Measurements", *J. Acoust. Soc. Am.*, 39 (1966), 151—168.

² Menzerath, P. and de Lacerda, A.: *Koartikulation, Steuerung und Lautabgrenzung* (Berlin 1933).

the production of an intervocalic consonant is greatly modified by the vowel context. In fact, the variability of the formant transitions contained in the initial vowel suggests that the speaker starts a gesture towards the final vowel while he is making the consonant gesture. It is as if the consonant gesture is superimposed on a diphthong movement.



$$S(x;t) = V(x) + K(t)[C(x) - V(x)]W_c(x)$$

$$V(x) = (1 - Q(t))\psi(x) + Q(t)V_2(x)$$

Fig. 1. Numerical Model of Coarticulation.

The block diagram of the lower part of Fig. 1 summarizes a strategy for the synthesis of time varying vocal tract shapes that reproduce coarticulation. This strategy has been implemented on a digital computer and tested against data collected from X-ray motion pictures of a human talker. Without going into details that have been published elsewhere,³ I should like only to draw attention to the separate

³ Öhman, S. E. G.: "Numerical Model of Coarticulation", *J. Acoust. Soc. Am.*, 41 (1967) 310-320

representations of the commands for the timing of the vowels and the consonants, and the commands specifying *which* consonants and vowels are to be synthesized. Target configurations for the initial vowel, the medial consonant, and the final vowel are fed to the model in the form of sets of numbers representing distances along the coordinate lines shown in the upper right corner of the figure. The consonant and

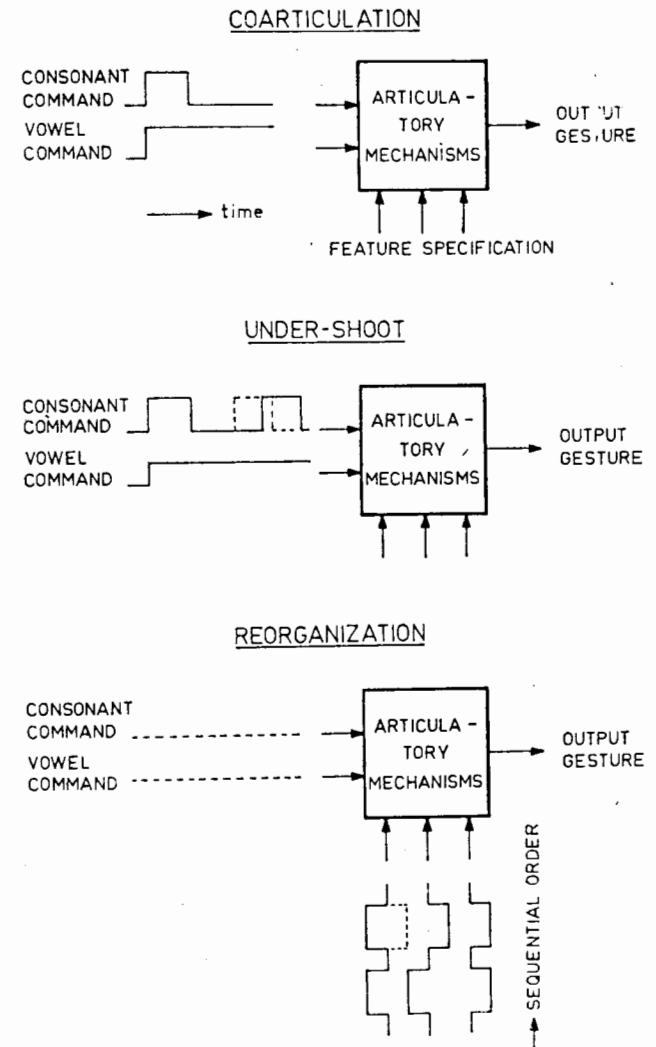


Fig. 2. Causes of Phonetic Variability.

vowel timing pulses shown on the left of the block diagram are then passed through the smoothing filters and enter the coarticulation model as one-dimensional signals marked K(t) and Q(t). A VCV gesture complex is then calculated according to the formulas shown below the block diagram. In this process the vowel diphthong ge-

sture is governed by $Q(t)$ and the superimposed consonant gesture is governed by $K(t)$. The calculation is done in such a way that, at the moment of consonantal closure, a residue of the underlying vowel gesture is always present in the vocal-tract configuration, so that the consonant becomes colored by the vowel environment. Hence variance is reproduced at the output while invariance is preserved at the input.

The articulatory mechanisms enclosed by dashed lines in Fig. 1 have been represented by a single "black box" in Fig. 2. Here, again, the timing commands are fed over two separate channels corresponding to consonants and vowels, and they are responsible together for the temporal integration of the syllable. The feature commands, on the other hand, specify the target configurations that the initial, medial, and final gestures of the syllable are aiming at.

In terms of this model the phonetic variance of phonological entities as observed at the acoustic level may be related to three types of physiological processes: *coarticulation*, *undershoot*, and *reorganization*.

We have already discussed coarticulation and this concept is illustrated again in the uppermost part of Fig. 2. Coarticulation results when the articulators are moving in response to distinct but temporally over-lapping commands.

Undershoot is indicated in the middle part of the figure by the difference in timing of the second consonant pulse. Undershoot thus results when an incomplete articulatory gesture is interrupted by a neural command that brings about the next gesture of the utterance. Examples of undershoot are found in the neutralization of vowels and consonants under increased rates of speech, as demonstrated by Lindblom.⁴ This type of variance may be accounted for quantitatively in terms of the numerical model discussed in connection with Fig. 1.

Reorganization, finally, is the result of a context dependent change of the feature specification, as illustrated in the bottom part of Fig. 2. This sort of effect is found, for example, in the devoicing of final voiced consonants in German and Russian, and in the quality alternations of vowels under vowel-harmony in a great many languages.

With respect to the underlying neural control there is thus an essential difference between the phonetic variance due to reorganization on the one hand and the variance typified by coarticulation and undershoot on the other hand. The sequence of feature specifications may be viewed as a phonological signal that modulates a phonetic carrier consisting of the standard timing pattern of the syllable. From this point of view reorganization is a perturbation of the modulating signal while coarticulation and undershoot are due to the structure of the carrier.

The model discussed so far has grown out of analyses of acoustic records and X-ray motion pictures. It is therefore of considerable interest to compare the picture of speech production summarized by the model with data obtained at the peripheral

⁴ Lindblom, B.: "Spectrographic Study of Vowel Reduction", *J. Acoust. Soc. Am.*, 35 (1963)a, pp. 1773-1781.

neural level by means of electromyographic methods. We shall here examine a few examples from a study in which thin concentric needle electrodes were used to record the motor unit activity from the facial muscles of a Swedish subject.

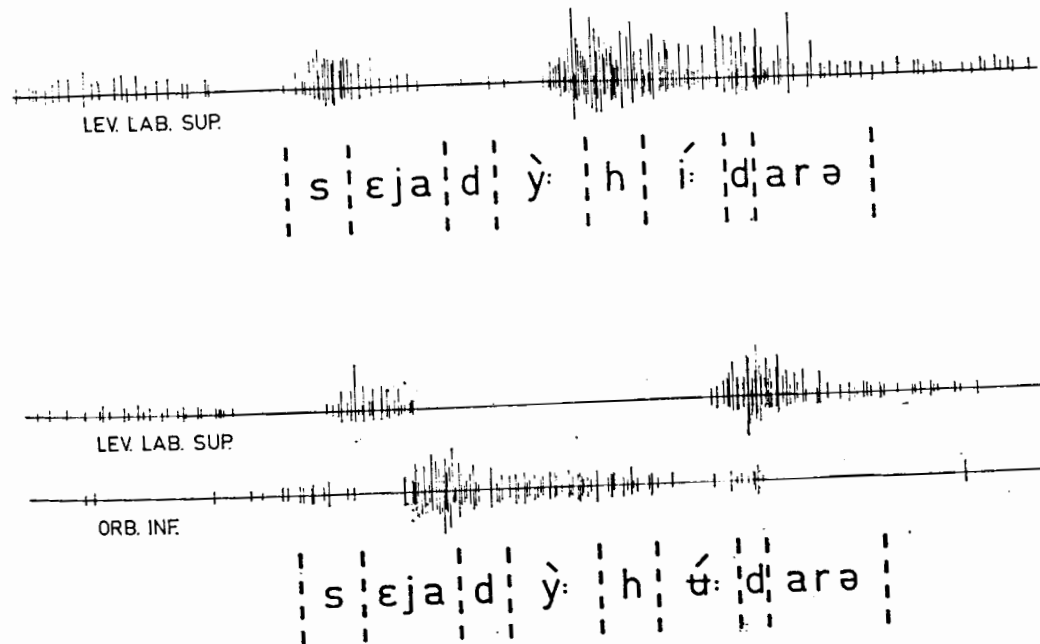


Fig. 3.

The top part of Fig. 3 shows a trace from a muscle that lifts the upper lip. In this record each motor unit potential is represented by a vertical line that indicates the amplitude of the spike. The VCV utterance $[y:hi:]$ is embedded in the frame $[seja d y:hi:dare]$.

Note that this muscle is tonically active *between* the utterances of the list that the subject read in the recording session. *During* the utterances, however, this background activity is depressed or enhanced in synchrony with the rounding and spreading gestures of the lips.

This behavior is typical and shows in speech the articulators assume a basic and apparently fixed posture on top of which excitatory and inhibitory motor commands are superposed as a modulating signal. This phonetic modulation of the basic speech posture occurs, of course, at a lower level than the phonological modulation of the syllable timing carrier, discussed earlier.

The phonetic modulation is also shown in the lower part of Fig. 3. The upper trace derives from the muscle just mentioned that lifts the upper lip, and the lower trace was picked up from the muscle that shortens and protrudes the lower lip. The utterance contains the VCV sequence $[y:hu:]$.

Note the reciprocal nature of these two signals. Whenever the lower trace shows

activity for rounding the upper trace is depressed, and vice versa. In this way the labial configuration as a whole is made to fluctuate about a constant average posture, so to speak.

Do we find invariance of motor commands at the peripheral neural level? Fig. 4 gives a negative answer to this question. The upper trace shows the motor unit activity

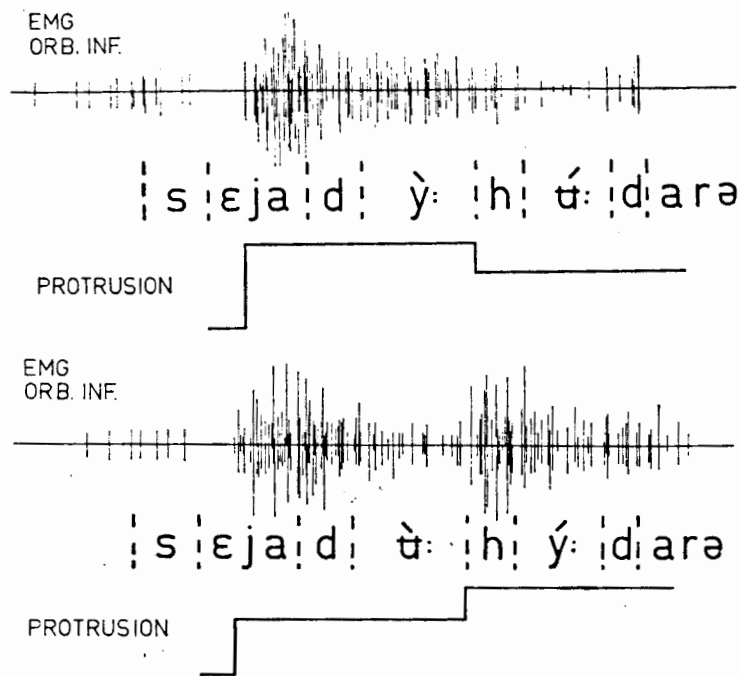


Fig. 4.

of the lower lip muscle referred to earlier. Here the VCV sequence $/y:h\epsilon/$ is embedded in the standard frame. In the utterance of the lower trace, recorded from the same lower lip muscle, the vowels of the VCV sequence have the opposite order, $/\epsilon:h\epsilon/$. The degree of protrusion for these two vowels is indicated schematically by the straight lines below the electromyographic records.

When the vowel $/\epsilon/$ is preceded by the less protruded vowels of the frame, a transitional overshoot appears at the beginning of the motor command. This overshoot is absent when $/\epsilon/$ follows the vowel $/y/$ which is more protruded. Hence the commands are not invariant. The steady state activity levels of these vowels seem to be less variable, however.

It is evident that the peripheral motor commands are calculated by the brain not only with respect to the constant effort necessary to maintain a target configuration, but also with respect to the variable effort needed to *move* the articulators from wherever they are to the desired target. It is quite likely that the last mentioned phase of

the calculation takes place at a rather peripheral level in solving feedback from the many sensory receptors in the oral region.

To sum up, a model of speech production—adequate for phonological purposes—should incorporate a sequence of stages of modulation as suggested in Fig. 5. As was emphasized by Dudley⁵ the acoustic sound pressure wave results from the modulation

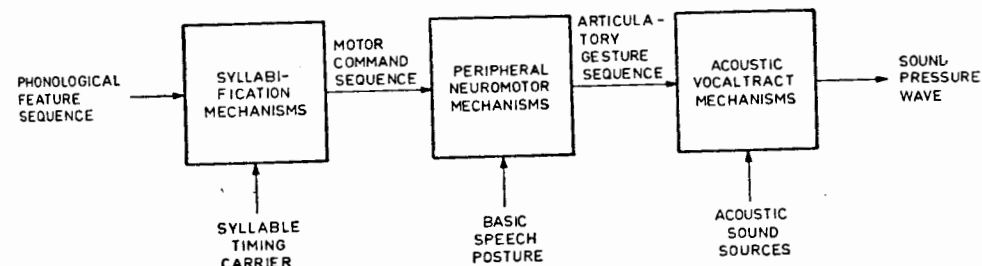


Fig. 5. Stages of Modulation in Speech Production.

of quasi-periodic or noisy sound sources by the relatively slow articulatory gestures. The latter gesture sequence comes about through the modulation of a basic speech posture by a sequence of excitatory and inhibitory peripheral commands. These commands, finally, may be regarded as the outputs of a set of more-central neural circuits through which a complex timing carrier is channeled by means of a certain phonological signal. This signal determines specific phonetic features of the various phases of the resulting syllables.

DISCUSSION

Fant:

1. Are the muscular movements programmed in advance only or is sensory feedback of basic importance?
2. Would it be possible to discover if unstressed vowels are modification of specific phonemes or independent allophones?
3. Why do you choose the term "undershoot" instead of reduction?

Lehiste:

I have been bothered for some time by the undershoot model of vowel reduction in unstressed syllables; it appears to me that the phenomenon can be language-bound and not predictable by the model. For example, post-tonic unstressed $/o/$ is realized as $[\partial]$ in Russian, but as $[u]$ in Bulgarian. The undershoot model does not explain the latter type of vowel reduction.

Tatham:

1. Give details of experimental procedure, please.
2. What relationship is there between the neural command arriving at the muscle fibre and the E.M.G. signal as measured, which is a result of the muscle fibre's reaction to the command?

⁵ Dudley, H.: "The Carrier Nature of Speech", *Bell System Techn. J.*, 19 (1940), 495—515.

EN MARGE DE QUELQUES PROBLÈMES PÉDOPHONÉTIQUES

KAREL OHNESORG

Si l'idée de M. Marcel Cohen, lancée en 1952, sur la fondation d'une chaire du langage enfantin dans chaque université¹ ne s'est pas encore réalisée, dans les dernières décennies, il a paru une longue série d'ouvrages nouveaux consacrés à l'étude du langage enfantin et qui ont déjà leur propre bibliographie, donnée par Werner Leopold et dont on prépare une édition élargie. Ces oeuvres font espérer, pour l'avenir, l'essor de la pédolinguistique et même de la pédophonétique comme disciplines indépendantes. Car les recherches linguistiques sur le langage des enfants dans divers pays fournissent, de nos jours, un matériel très riche et qui pourra servir de base à des études comparatives du développement du système phonologique et de l'évolution phonétique du langage enfantin. Et, ce sera le rôle de la pédophonétique de suivre et d'expliquer, comment les réalisations communes tout d'abord à tous les enfants se différencient et s'accommodent à celles de la langue de l'entourage.

Cet accommodement se réalise tout d'abord à l'aide de l'ouïe. La distinction des consonnes *b-m*, *t-k* ne peut être expliquée que par leur perception acoustique: l'abaissement du voile du palais pour la formation de *m* et la position postérieure de la langue pour le *k* échappent à la perception visuelle. C'est pourquoi les enfants des peuples divers (même les enfants chinois) interchangent souvent, au début, ces consonnes, et cela quelquefois même à l'époque où ils ont déjà acquis leur réalisation correcte.

On pourrait supposer les mêmes causes pour la confusion élémentaire — et si fréquente — commune à tous les enfants du monde entier, des consonnes *l* et *r*. Mais, ici c'est aussi l'habileté articuloire qui entre en jeu et qui se développe graduellement, comme suite des justes perceptions acoustiques.

Dans le domaine de la réalisation des phonèmes, dans le langage enfantin, la concordance des faits constatés par divers auteurs est plus que surprenante. Pour démontrer qu'on peut découvrir, dans la formation phonétique du langage enfantin, des règles générales ou même certaines lois, insistons, dans notre court rapport, sur quelques-unes d'entre elles. Car on peut, de nos jours, confirmer et enrichir la liste des consta-

¹ Marcel Cohen, *Sur l'étude du langage enfantin*, Paris 1952, p. 181: „Une science bien organisée devrait comprendre des chaires de langage enfantin...“

tations faites déjà dans la première moitié de notre siècle par M. Roman Jakobson: Les enfants prolongent la voyelle pour compenser la consonne suivante qu'ils omettent; les sons caractéristiques pour une certaine langue se développent et se fixent dans la langue des enfants le plus tard (p.e. les voyelles nasales ou labialisées). — Les enfants substituent tout d'abord les occlusives aux fricatives homorganes, ils négligent la différence entre les consonnes sourdes et sonores (ou tendues et relâchées) avec la prédilection pour les sourdes, les consonnes sourdes se fixent dans leur parole plus tôt que les sonores (même avec la distance de plusieurs mois, comme c'est souvent le cas de la paire *f-v*), ils confondent les phonèmes *s* et *š* dans une seule réalisation (*s* palatalisé), ils palatalisent souvent les consonnes (et cela même dans les langues où la palatalisation n'a point lieu, parfois ils tronquent les mots ou les groupes rythmiques contenant plusieurs syllabes).

Ce dernier phénomène a été signalé par beaucoup d'auteurs. Mais, ils en donnent des explications différentes. Les uns sont persuadés que l'enfant perçoit tout d'abord et rend ensuite les syllabes accentuées. Mais, on peut trouver, dans le vocabulaire des enfants, aussi des formes où la fin du mot est conservée, tandis que les syllabes (ou les consonnes) initiales tombent quoiqu'elles soient accentuées, comme c'est par exemple en tchèque. C'est pourquoi d'autres auteurs expliquent ces formes raccourcies comme suite de l'écholalie, si fréquente dans les débuts de la parole de tous les enfants. Par une étude détaillée de ce phénomène, nous sommes persuadés que les deux facteurs entrent ici en jeu: tout d'abord, les formes raccourcies ont leur origine dans l'écholalie (de la même manière que les premières réactions vocales aux questions consistent en la répétition des dernières syllabes de la question). Mais, plus le langage de l'enfant s'intellectualise, plus l'enfant retient les syllabes accentuées ou le radical des mots. Ce qui prouve que O. Jespersen, dans ses considérations sur les mots hypocoristiques,² avait supposé à juste titre que les formes raccourcies où le commencement du mot est conservé, ont été créées par les adultes, tandis que celles où la fin du mot est maintenue, ont leur origine dans le langage des enfants.

De cette constatation, il en résulte une conclusion importante: dorénavant, les plus détaillées et les plus soigneuses analyses linguistiques du langage enfantin ne sauront se passer d'une observation psychologique simultanée. La collaboration des linguistes et des psychologues pourra ensuite contribuer à l'explication de nombreux faits, soit qu'il s'agisse du rôle de l'intonation aux débuts du langage enfantin, soit que l'on insiste sur la fonction de l'accent, soit que l'on veuille étudier l'influence du milieu ambiant dans sa complexité, expliquer la prédilection des enfants pour le rythme et la rime ou leur penchant pour les expressions insolites par leur nature acoustique.

Outre cela, quelques déviations articulatoires des enfants, causées par l'émotion, sont analogues à celles qu'on peut saisir dans le langage empreint d'emphase des

² Otto Jespersen, *Language*, London 1922, IX, 7 (Stumpwords).

adultes et, souvent elles sont les mêmes (p.e. la désonorisation des consonnes, quelques métathèses). L'analyse de ces faits pourra faire découvrir et éclairer les tendances latentes de l'évolution de la langue qui de nos jours est fixée et liée par la langue écrite.

Cependant, le domaine de l'analyse de la langue des enfants devient de plus en plus vaste, les problèmes augmentent, les matériaux qui faciliteront des études comparatives s'accroissent. L'investigation de l'évolution et de la fixation des articulations enfantines, à l'aide des appareils modernes, et en même temps, la collaboration des linguistes et des psychologues, permettront de constater la nature acoustique et articulatoire du langage enfantin et, en même temps, la dépendance de la parole des facteurs psychiques. Ce qui pourra contribuer à la solution du problème proposé par Mme Slama-Cazacu comme un sujet de recherches internationales:³ quelle est la voie de l'évolution des sons biologiques par préphonèmes aux phonèmes, dans le langage enfantin. D'ailleurs, on ne peut pas douter que les résultats ne soient pas sans importance même pour la phonétique et la linguistique générales. Toujours est-il, que la constitution de la pédophonétique et de la pédolinguistique comme disciplines indépendantes ne saura désormais être contestée.

DISCUSSION

De Vriendt:

L'auteur a souligné la nécessité d'une observation psychologique simultanée pour toutes les études de pédophonétique. Un examen sociologique parallèle nous semble tout aussi indispensable (voir Travaux de Bernstein à Londres, Enquête parmi la population noire de Washington etc.).

Cohen:

On ne peut que se réjouir de voir se continuer les utiles travaux du Professeur Ohnesorg sur le langage enfantin et de l'entendre exposer ses projets pour l'avenir en une matière aussi importante, qui devrait être représentée dans les universités de tous pays.

³ Tatiana Slama-Cazacu, *Quelques remarques psycholinguistiques sur l'apprentissage de la langue par l'enfant*. — *Langage et comportement* 1 (1965) 181—193; *Cahiers de linguistique théorique et appliquée* 3 (1966) 171—179 (București).

PHONEMIC SYNCRETISM AND THE PROCESS OF COMMUNICATION

ZDENĚK F. OLIVERIUS

The problem of phonemic syncretism, neutralisation, overlapping of phonemes etc. has been widely discussed by linguists of different schools, of whom I might mention primarily N. S. Trubetzkoy (8), A. Martinet (6), L. Hjelmslev (3), Z. S. Harris (2), A. A. Reformatskij (7) and R. I. Avanesov (1).

The failure of many attempts at automatic detection makes it advisable to revise the concept of minimal linguistic units.

In this contribution I would like to draw attention to the interrelationship of certain features of languages as natural communicative systems and the concept of phonemes as minimal units of the expression level.

The minimal elements of the expression level are usually defined with the tacit presumption that a one-to-one correspondence can be achieved solely on the basis of an analysis of inherent qualities of individual segments. It is assumed that every single segment counts.

This would be the case in a language with zero redundancy (where all possible sequences of segments constitute messages): in such a case only the acoustic image of every single segment would have to be taken into account (not the acoustic images of neighbouring segments); one acoustic image could have — in case of zero redundancy — one and only one function.

The point of departure of this contribution is the obvious fact of greater than zero redundancy in actual communicative systems (with the necessary implication of a) absolute limitations on sequence of segments and b) variations in relative frequency of different sequences).

Stringent limitations as to which segments may occur in utterances permit us to differentiate and keep apart two acoustically identical segments; the environment (sufficiently wide, in some cases extending to the limits of the utterance) takes over the role of the inherent acoustic correlates of distinctive features. The phonemes will consequently be understood not as a bundle of distinctive features but as a class of elements united by one function (identifiable by the inherent features of the given elements plus its environment, i.e. the acoustic features of surrounding segments). Under such circumstances complete overlapping of phonemes is possible. It can be dealt with as a (resolvable or irresolvable) syncretism on the phonemic level.

Two acoustically identical segments [t] and [t] in words *pod* and *pot* or two other

identical segments [ə] and [ə] in words мальчиком and мальчиком never appear isolated in normal communicative situations. Segments can be ascribed to a phoneme (as members of a class of elements to the class in question) only within normal unambiguous utterance, as for example:

- (1) Его богатство, знатный род и связи давали ему большой вес в губерниях, где находилось его имение. (Пушкин, Дубровский)
- (2) Не шумели, не гадали и оратору глядели прямо в рот. (Д. Бедный, Друг надежный)
- (3) Что за короткость с пустым, глухим мальчиком? (А. Островский, Бесприданница)
- (4) Дима об этом еще ничего не сказал мальчиком.

In the quoted pairs of utterances (1—2 and 3—4) the acoustically identical elements are in different positions (if sufficiently wide positions are drawn in) and permit the full application of Trubetzkoy's I and III rules for identifying phonemes (8,42—44), and also the application of the strictly distributional criteria of Z. S. Harris (2,61).

If we impose the (in linguistics) usual limitation on the analysis of the segments in question and take into account only the inherent acoustic qualities of the segment and a very narrow position, we cannot solve the syncretism and we have to admit the existence of an archiphoneme (i.e. of an element with a limited distinctive power, which is evidently contradictory to the accepted presumption that expression is by necessity built up of discrete elements (which must be either in opposition to each other or be identical—tertium non datur). cf. B. Mandelbrot (5).

But such a limitation—even if it is usual in linguistic analysis—is not imposed on the analysis done by native speakers in the process of decoding utterances containing phonemic syncretisms like род — рот, мальчиком — мальчиком. Otherwise decoding would be impossible and the interlocutor would demand further explanation. In decoding syncretisms like род — рот, мальчиком — мальчиком the listener takes into account the whole sequence of segments before and/or after the element in question to the limits of the whole utterance (or at least certain pieces of the sequence like ... знатный ... х ... глядели в ... от ... с ... х ... не сказал ...

The structured message is encoded into a sequence of discrete elements (phonemes). The discrete character of signs and figurae (cf. L. Hjelmslev [3]) is equivalent to the fact that a message can be repeated without distortion and disfiguration. If signs and figurae had the character of continuous elements they could be only imitated (and not repeated), which would of necessity imply a steadily growing disfiguration of a message at each repetition. Communication would be practically impossible under such conditions (cf. M. Mandelbrot [5]).

Discrete elements (phonemes) are, in turn, transformed (via nerve impulses) into continuous muscular activity, then into a continuous sound-wave. The sound-wave

reaches the hearing apparatus of the receiver where it is decoded as a sequence of discrete elements again.

At the different stages of encoding and decoding a message we are dealing with a transposition from one code to another. While in normal communication elements of the content level remain identical from the first to the last stage of the process of communication, the elements of the expression level (if I may use the term "expression level" in this rather wide meaning) change from nerve impulses to muscular activity and different shapes of the sound-wave etc. Segments of the successive different subcodes (e.g. the muscular activity, the sound-wave) can be related to the same phoneme. So the discrete element, the phoneme, can be said to correspond with segments of different successive continuous expressions.

The usual definitions of phonemes drawing attention only to one or two stages of the process of communication (the muscular activity and/or the shape of the sound-wave) are not adequate enough.

We can assume that some elements (phonemes) are kept apart at the stage of nerve impulses and neutralised at the stage of muscular activity and sound-wave (many examples could be drawn from Slavonic languages, e.g. Russian and Czech). It seems fruitful to adopt the stimulating conception of A. V. Isačenko, who suggests the presupposition that phonemes are encoded as a complete set of instructions (4,204).

The concept of phonemes as minimal discrete units of the expression level should take into account all stages of the process of communication with different continuous realizations.

Simple and adequate description of the expression level of a language can be achieved by means of only two types of elements: discrete phonemes and continuous allophones. Archiphonemes, morphonemes etc. are not indispensable.

BIBLIOGRAPHY

1. P. И., Аванесов, *Фонетика современного русского литературного языка*, Москва 1956.
2. Z. S. Harris, *Methods in Structural Linguistics*, Chicago 1963.
3. L. Hjelmslev, *Omkring sprogteoriens grundlæggelse*, København 1943.
4. A. V. Isačenko, Foném a jeho signálový korelát, *Slovo a slovesnost* 27 (1966), 193—205.
5. B. Mandelbrot, Structure formelle de textes et communication, *Word* 10 (1954), 1—27.
6. A. Martinet, Neutralisation et archiphonème. *TCLP* 6, 1936
7. А. А. Реформатский, О соотношении фонетики и грамматики (морфологии), *Вопросы грамматического строя*, Москва, 1955.
8. N. S. Trubetzkoy, *Die Grundzüge der Phonologie*, *TCLP* 7, 1939.

SOME NOTES ON VOWELS AND GEMINATED CONSONANTS IN CONTEMPORARY ITALIAN

JAROSLAV ONDRÁČEK

Before trying to point out some characteristic features of the Italian phonological system as regards vowels and geminated consonants, I should like to say that the type of pronunciation referred to is the norm described by a number of Italian linguists (for instance, A. Camilli) and recently adopted in the *Dizionario Garzanti della lingua italiana* (published in Milan in 1965). This sort of "received pronunciation" is also the basis for the majority of phonetic (or phonematic) transcriptions, for example by Piero Fiorelli, whose texts alternate with those of Marguerite Chapallaz in the journal *Le maître phonétique*. The lack of consistency that we sometimes find in such transcriptions is not present in Fiorelli's orthoepic spelling (used in his book *Corso di pronunzia italiana*, Padova 1965), which is an interesting compromise between phonetic transcription and normal spelling.

It is known that most Italian words are stressed on the last syllable but one. If we take this type of word as a model for the contact between vowel and consonant, we can see that the possible phonological realization is determined by certain relationships between the sounds. Our selection could contain words with intensified consonants, which are really a characteristic feature of Italian pronunciation, for example: *tappo*, *fatto*, *tacco*; *tuffo*, *passa*, *fascia*, *pazza*, *faccia*, *mamma*, *panna*, *bagna*; *palla*, *paglia*, *narra*. The geminated consonant here reaches the highest degree of intensity, whereas the duration of the vowel is at the lowest point. Starting from this extreme case, we arrive through various consonant groups to the opposite combination, where the stressed vowel is lengthened and the consonant is unintensified: *tipo*, *fato*, *fico*; *tifo*, *naso*, *tace*, *lame*, *mano*, *palo*, *caro*.

Of course, there are other factors to help distinguish between words: the voiced or voiceless quality of consonants, and the closed or open quality of vowels. That means that intensified voiceless consonants, except *ss* and *sc(i)*, are also voiced, and the vowels E, O can have a closed sound or an open sound: *ebbe*, *cadde*, *lèggo*; *bèveve*; *mèzzèze*, *lògge*.

The value of distinctive features changes according to the requirements of the meaning and according to the character of the sounds themselves and their position. For example, the consonants *sc(i)*, *z*, *z*, *gn*, *gl(i)* in the words *fascia*, *pazza*, *gazza*, *bagna*, *paglia* are intensified (*z* and *z* are doubled in spelling) and the preceding vowels are more or less short. The spelling and unintensified pronunciation of the voiceless

z in words like *Venezia*, *vizio* (Fiorelli admits this pronunciation) shows the possibility of differentiating between *vizi* and *vizzi*, *tizi* and *tizzi*, *spazi* and *spazzi*, but the majority of Italian linguists (among them Camilli, Fiorelli, Migliorini) usually make no difference between the pronunciation of z and zz. I think that a lengthened vowel before one z can be heard quite clearly if we compare the word *avvezzo* with the words *spazi*, *grazie*, *orifizio*, *Fabrizio*. And spelling does not avoid the doubling of z, common in other consonants; I am referring to pairs like *battere*—*abbattere*, to which we may add the pairs *zitto*—*azzittiere*, *zòppo*—*azzoppare*, *zèro*—*azzerare*.

As far as vowels are concerned, the scale of quality goes from open A to open or closed E and O and finishes with closed I and U; the pronunciation of the last two sounds can also be semiconsonantal: *pali*, *èri*, *pìle*, *ìeri*, *òro*, *óre*, *pure*, *nuòre*.

Now let us look at the contact between vowels. In the so-called rising diphthongs (*ìeri*, *nuòra*) the semiconsonantal element is relatively stable, but the other order (vowel—semiconsonant), which occurs in the so-called falling diphthongs, is on the other hand characterized by the possibility of forming hiatus. Words like *mai*, *poi*, *sei* sometimes have (especially in singing) a bisyllabic pronunciation because the semiconsonant becomes a vowel here. When we add that neither the combination *i* or *u* plus a stressed vowel assures the realization of a diphthong (*viòla*, *continüare*), it can be seen that unlike other languages (for example, Czech, where the glottal stop or a transitional sound may be used) Italian forms hiatus very frequently; it appears not only in the middle of a word but also between words in a sentence. And we may note that such contacts lead quite often to the formation of new diphthongs or to the omission of the final consonant. Fiorelli illustrates this with the following examples: *parrà oscuro* (hiatus or diphthong), *pare ovvio* (diphthong or hiatus), *pare oscuro* (diphthong), *quell'affare* (elision), *par oscuro* (troncamento).

We have mentioned the word *mai* and its possible bisyllabic pronunciation with two real vowels. I should say that we are not very far from groups in which a stressed (and, consequently, lengthened) vowel is followed by another, unstressed vowel, both being pronounced; with "one impulse of the breath" (I. C. Ward). Let us remember for a moment the English diphthongs and try to find some Italian words that might serve as counterparts. For example: try [*trai*]—*tràe*, now [*nau*]—*ciàco*, dare [*dæə*]—*dèa*. Considering the quality of the Italian final vowels and the character of the English mixed vowel, the comparison seems to be useful. Other examples may tell us something, too: mere [*miə*]—*mìa*, tour [*tuə*]—*tù* etc. The difference between *ma-i* and *mi-a* lies in the direction from one element of the group to the other. So one impulse is able to form a diphthong both in the word *mai* and in the word *mia*. Some Italian phoneticians are aware of this possibility and draw attention to it in their transcriptions (for instance, P. Fiorelli in *Le maître phonétique*) by indicating clearly the semivowels (*maï*, *miä*); the first element in such vowel groups is naturally more prominent than the other, so they might be interpreted as falling diphthongs.

When we were talking about the contact between vowel and consonant, we referred to the lengthening of the stressed vowel. Its duration, however, has nothing to do

with the so-called quantity which exists in other languages (in Czech or in Finnish, for example) and is independent of stress. While Czech, a language with the stress always on the first syllable, makes use of short and long vowels without regard to accentuation, the length of Italian vowels is only potential and is determined by their stress and position in the word. The "quantity" functions here as a prosodic feature.

DISCUSSION

Wittock

L'idée de M. l'abbé Falchun, prononcée tout à l'heure, pourrait être très utile pour l'étude de l'évolution des diphtongues dans les langues *romanes*, par ex. (Mentionnons, parmi autres, l'espagnol avec son influence des palatales sur l'évolution de la voyelle diphtonguée précédente). Nous croyons que cette idée pourrait servir de base pour un système entier de recherches concernant le développement des langues nommées.

ELECTRICAL GLOTTOGRAPHY

J. ONDRÁČKOVÁ*, J. LINDQVIST,** G. FANT**

Our study was concerned with a glottograph of the Fabre type belonging to the Laboratory of Phonetics in Prague. The principle of the Fabre glottograph is that the movements of the vocal folds are picked up as an electrical resistance change between two electrodes placed on each side of the thyroid cartilage.

Fig. 1 shows the signal response of the Fabre glottogram when a relay suddenly increases a resistance in parallel with the electrodes. This temporary resistance increase is associated with an upward deflection of the glottogram curve. Two different lower limiting frequencies are used in the glottograph depending on which purpose it is used for. The direct output of the glottograph with a time constant of 4 sec, displays a rather clean rectangular response. The built-in amplifier curve shows a waveform distortion typical of the 25 msec time constant. As seen from B this distortion is not serious in studies of stationary phonation, providing the duration of the pitch cycle is not excessively long.

Once the vocal cords have opened in their full length a further separation does not affect the resistance much. Thus the flat top of the cycle. On the other hand, once the vocal cords, when closing, have made contact at some point there will follow a further rapid resistance decrease because of the contact progressing horizontally and vertically, thus offering an increased area of conductance for the glottograph current. The following phase from maximum closure to opening is generally more gradual at a low voice pitch. This is to be expected in view of the axial component of the vocal cord movement which during the closed period displaces the area of contact upwards whilst this area is decreasing. One can often identify a point of contact break as indicated by a sudden change of the slope of the glottogram curve.

The outstanding feature of the Fabre glottogram is the extremely rapid resistance fall when the cords reach contact in the closing phase. It is therefore well suited for measuring the fundamental frequency in speech.

The Sonesson photo-electrical glottography provides a measure of the cross-sectional area at the level of maximum narrowing in the glottis. To a first order of

* Laboratory of Phonetics, Československá akademie věd, Prague.

** Speech Transmission Laboratory, Department of Speech Communication, Royal Institute of Technology, Stockholm.

approximation this area as a function of time is proportional to the flow of air through the glottis, as theoretically predicted by van den Berg et al (1957). It is also known that inverse filtering of the current from a pressure microphone with extended low frequency response provides a measure of this glottal flow. We can in Fig. 2 compare these two methods with the Fabre glottogram.

In order to bring out the synchrony with the acoustic flow in C and D, we have corrected for a 1 msec latency in the inverse filter, as indicated by the displacement of the time scales. We can now observe that when the electrical glottogram in the ascending opening phase has reached the saturated flat level, the glottal pulse flow has already reached one-half of its peak value. This should be the instant when the glottal slit has reached a state of maximal length while the cords continue to separate. At the instant when the electrical glottogram signals the first contact of the cords, the air pulse is almost but not quite terminated.

In Fig. 3 there is displayed the oscillogram of three sentences and the direct output of the Fabre glottograph below. The time scale is compressed in this figure. Well ahead of the phonation the resistance keeps falling by a magnitude greater than the amplitude of periodic excursions within the voiced portion. To what extent this is a result of overall changes in the spatial distribution of muscle mass and to what extent it reflects the successive closing of the arythenoids is hard to say.

In A of Fig. 3 we see the typical excursion of the base line upward during the voiced occlusion of /b/ whilst the voicing ripple successively decays in amplitude. In less stressed articulations of voiced stops the glottograph oscillations retain rather constant amplitude during the voiced occlusion. In the unvoiced stop gap of /k/ there is a similar but more pronounced increase of the resistance signifying a corresponding opening movement. The extent of this never reaches the value of rest before the sentence.

C finally shows that a glottal stop inserted to function as a boundary marker between two morphemes, one ending with a vowel and the other starting with a vowel. The resistance retains a value even lower than in the preceding and following voiced segments. This is natural in view of the cords being tightly pressed together.

Preliminary studies of the glottogram mean resistance as a correlate of the distinctions between stressed and unstressed syllables and as a factor influenced by the Swedish word accent 1 (acute) and accent 2 (grave) have been undertaken but these are not conclusive enough to deserve a detailed discussion. There is some indication that a stressed vowel has a lower resistance than an unstressed vowel but this is not quite consistent.

The Fabre glottograph, when adjusted to have a long time constant, is a valuable instrument for indirect studies of the voice mechanism as well as of laryngeal articulations. It is recommended as a supplement to inverse filtering or optical glottography. The typical feature of an extremely well developed discontinuity in the closing phase could be made use of in speech analysis as a reference for time domain voice frequency analysis.

REFERENCES

- van den Berg, Jw., Zantema, J. T., and Doornenbal, Jr., P.: "On the Air Resistance and the Bernoulli Effect of the Human Larynx", *J. Acoust. Soc. Am.* 29 (1957), 626—631.
- Sonesson, B.: *On the Anatomy and Vibratory Pattern of the Human Vocal Folds, Acta Oto-Laryngologica*, Suppl. 156 (Lund 1960).
- Further references see in *STL-QPSR* 4/1966 (Fant, G., Ondráčková, J., Lindqvist, J. and Sonesson, B.)

DISCUSSION

Fischer-Jørgensen:

Thanks to a visit by dr. Ondráčková to Copenhagen we have also had the opportunity to work with Fabre's glottograph and we got very similar results for consonants. Difference according to vowel quality intonation and stress seem, however, to indicate that the curves may be influenced by other factors than the degree of opening of the glottis, e.g. the gross movements of the larynx. We should suggest experiments with smaller electrodes having a well defined placement (see Annual Report of the Phonetic Institute of Copenhagen 1967).

Gsell:

I would be interested in technical data about the device of the glottograph you built. In Grenoble we tried to construct the classical glottograph as described by Fabre. The results we obtained with this basic device were close to those obtained with an intensity-meter on voiced and unvoiced parts of speech.

Ondráčková, Lindqvist, Fant:

Ad Gsell: The technical design is not much different from that used by Fabre and Husson, see also *STL-QPSR* 4/1966. In our experience the glottograph trace differs substantially from that of an intensity record. The amplitude of the glottal periods are almost constant as long as the phonatory mode is constant. The characteristic feature of the glottographic wave shape is the rapid step signalling the instant when the cords make contact which can be adopted as a pitch period marker.

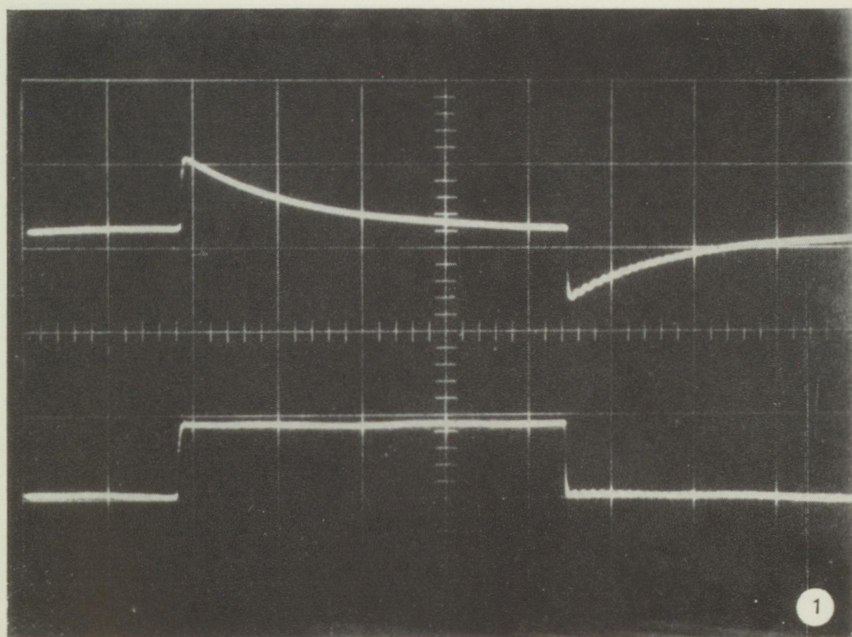
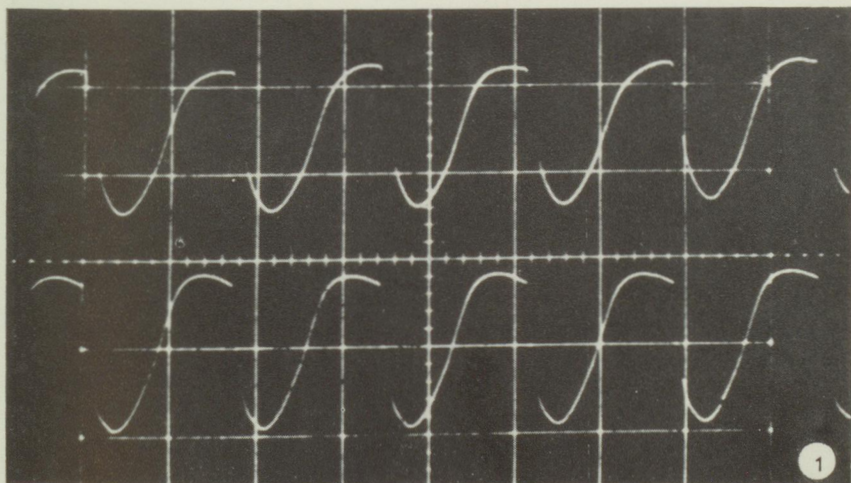


Fig. 1. A. Response of the glottograph to a rectangularly shaped resistance increase. The direct output with a long time constant preserves the waveform.



1. B. Typical glottogram, male subject, long time constant above, short time constant below synchronously recorded. ($\tau = 4$ sec; $\tau = 0,025$ sec)

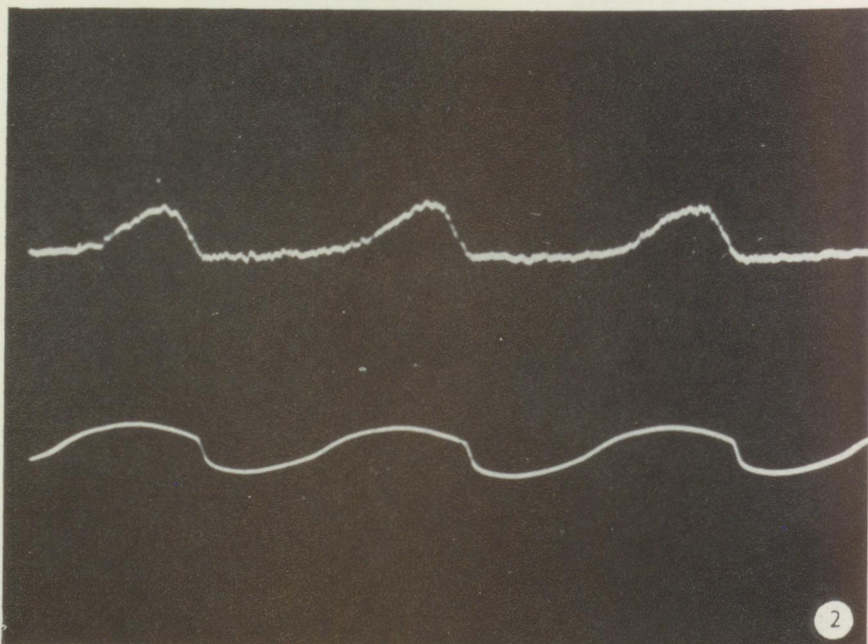
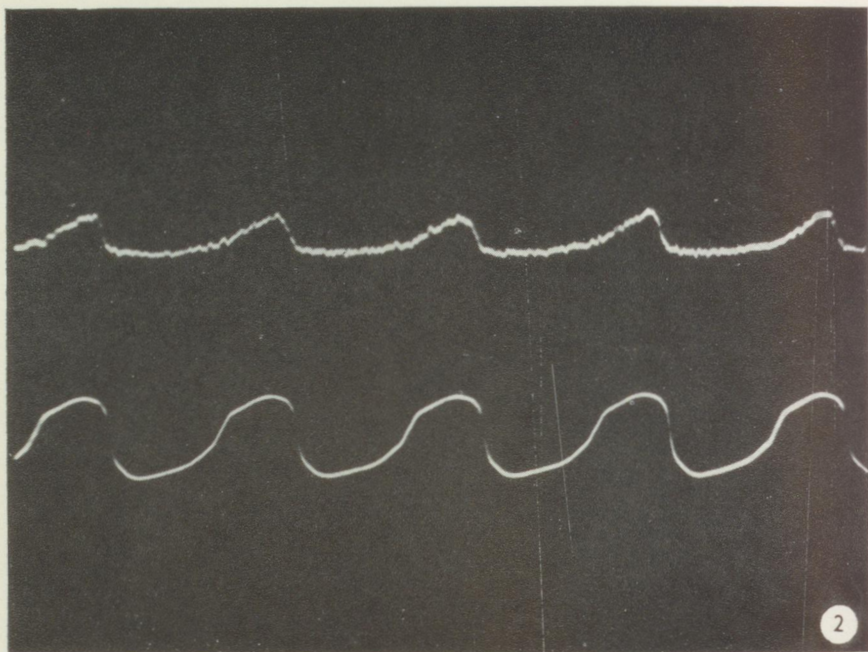


Fig. 2. A, B. Synchronous recordings of optical (above) and electrical glottogram (below), speaker (GF).

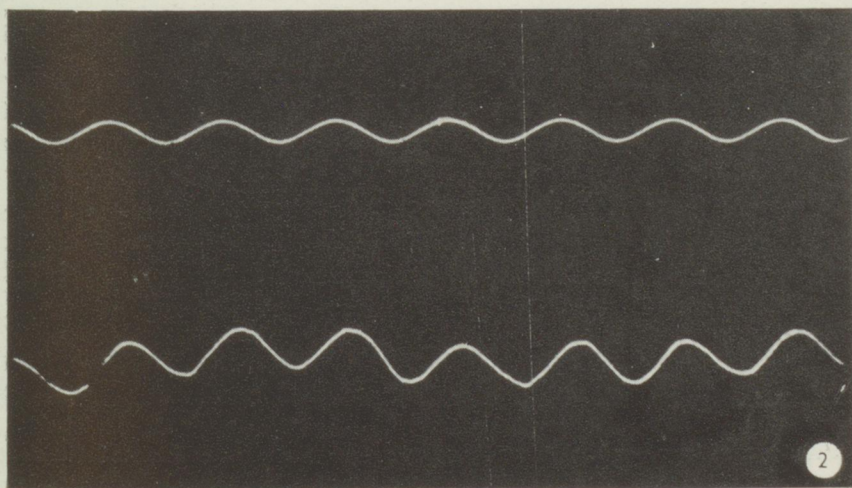
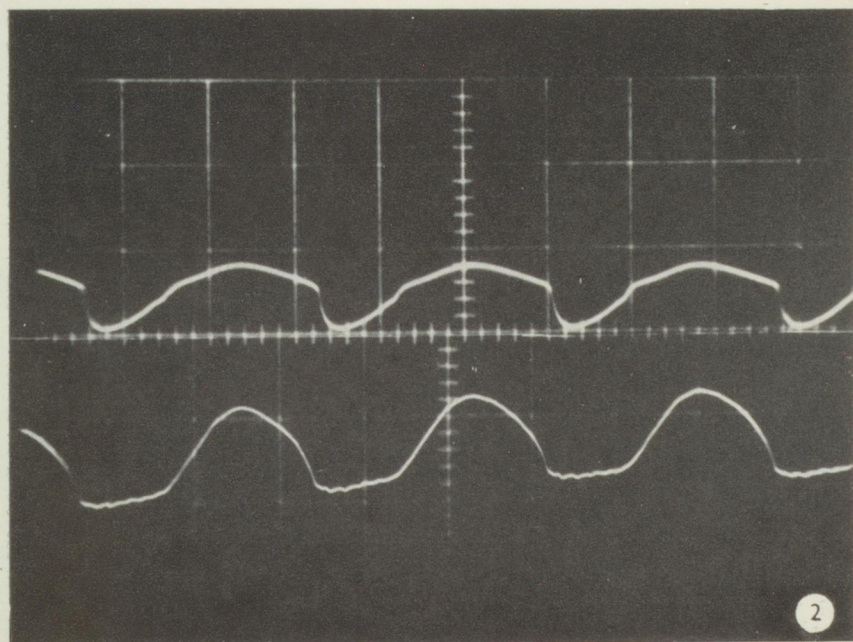


Fig. 2. C, D. Electrical (above) and inverse filter glottal flow (below), subject (JS).

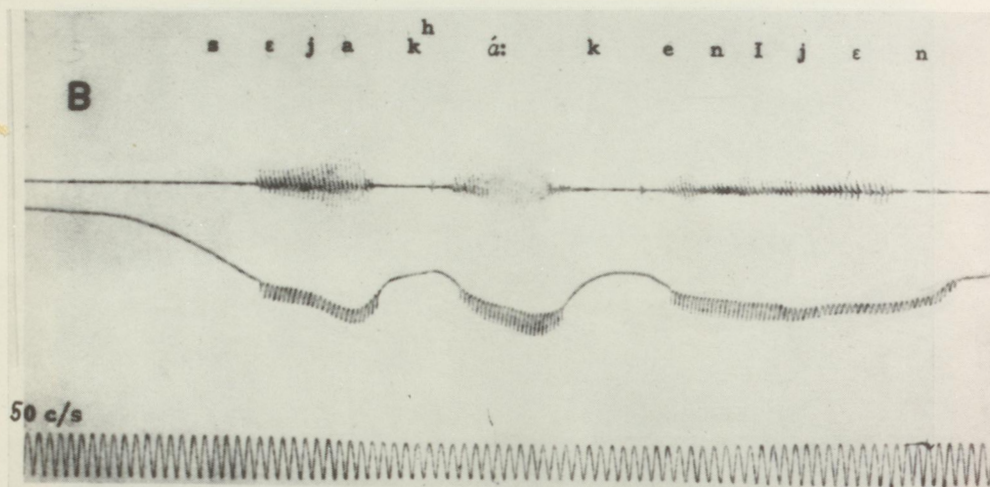
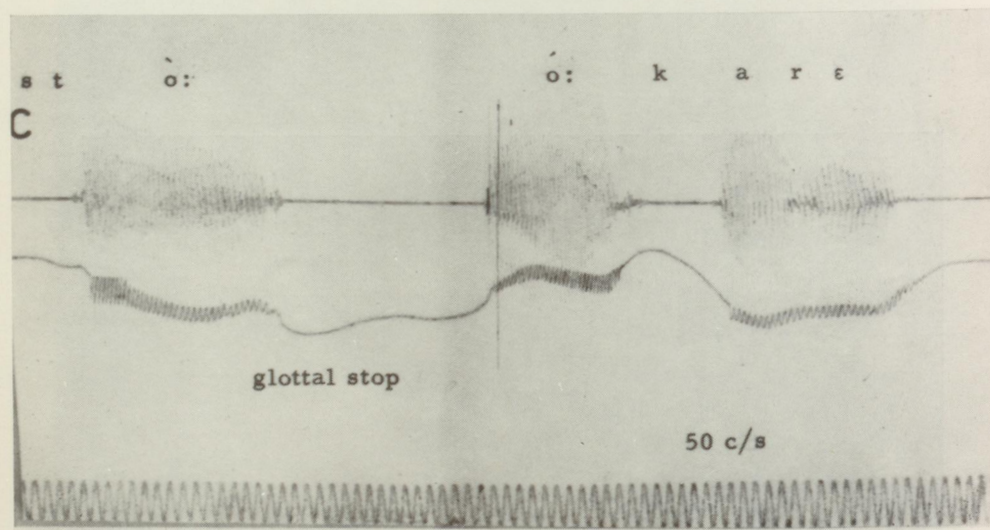


Fig. 3. B illustrating the accent I nonsense word [ká:ken] within a carrier phrase.



3. C. Glottal stop marking a word boundary shows up as an extreme low resistance in the electrical glottogram. The Mingograph channel displaying the output of the electrical glottograph lags the speech oscillogram trace by 15 msec as indicated by the vertical line. This instrumental factor affects (A), (B), and (C).

PERCEPTION OF QUALITY AGAINST QUANTITY

MASAO ONISHI*

Which is more important, quality or quantity of sound? The perception of "long sounds" and "short sounds" will be easy to understand as they are common and practical. But, is the length of a sound absolute? No! For instance, the English dictionary reads "do" as [du:], "pass" as [pa:s], and "me" as [mi:]. They are all long vowels but in actual speech they are often pronounced as short or semi-short like [du] or [duː] [pas] or [paːs] [mi] or [miː]. It is proved by the oscillograph test that the physical absolute quantity of so-called vowels differs according to the occasion and situation. I wonder therefore if it is right or not that the quantity of sounds which is changeable is fixed in dictionaries, giving us the idea that the meaning of a word is not changeable, either. Conversely short vowels can be prolonged and changed to long vowels. For instance, "good" [gud] is often pronounced as [gu:d]. It is very common that "little" [lɪtl] and "very" [veri] are prolonged like [lɪ:tl] and [ve:ri]. The pronunciation of words is always influenced by *intonation* and *prominence* and it is very often changed in *length*.

From the above view point, I do not like to agree with the notations wherein quality may be underestimated because of the idea of the fixed quantity of sounds. In general, the teaching of languages must follow the auditory tracing of *Lautgefühl* and it is desired that the notations show the sensitive difference of the quality of sounds in an appropriate way. I believe that the *IPA-system* is extremely adequate in this point. I do hope that this system should be used in general English dictionaries and in English teaching of the world, drawing distinction particularly *ɪ* from *i*, *e* from *ɛ*, *ɔ* from *o*, and *ɜ* from *ə*, as they are luckily adopted by *Le maître phonétique*. I believe that this theory can be applied effectively to other languages than English and that it is an important factor to promote the study of Phonetic Sciences themselves.

As a proof of the above, I will explain the characteristics of the Japanese language and the important question of quality against quantity of sounds. It is commonly believed that the Japanese language has long and short consonants and vowels. It seems that the Japanese themselves have been brought up with this pre-conception.

* Hosei University, Tokyo.

However, this is not correct from the current scientific standpoint and furthermore, it may be a wrong course which prevents the Japanese language from phonetic development. The alphabet peculiar to Japanese is the so-called "kana" a kind of syllabary. Each of the syllabaries has no rule or usage in record as to whether it is short or long. The rule has not been settled yet and the sound is changeable from "long" to "short" and vice versa. In short, Japanese syllabaries have no physical absoluteness but they are pronounced on the basis of psychological feeling. Strictly speaking from the acoustical standpoint,—setting mechanical sounds apart—any human voice, not limited to Japanese, which is based on breath movement will never have a pure long sound. Every sound changes its quality in accordance with the *duration of time*.

Particularly in Japanese,—which has the characteristics of a tone-language or musical accent language—when a sound which is considered "short" is prolonged, tonal transition takes place without exception. The Japanese language is generally called a musical accented language against stress accented languages such as English, French and German by conventional students of the Japanese accent including foreigners. It appears that this theory has developed into a kind of faith among the researchers of linguistics. In other words, they assume that a Japanese word composed of a certain number of syllables is pronounced so that a syllable (or syllables) of the word is (or are) pitched higher or lower than the other syllables. They try to summarize, patternize and stabilize the pitch form and make a perception out of it. On the other hand, however, all the Japanese people except a few scholars having interest in the discussion of accent—including, of course, teachers of primary schools—are almost ignorant of the word-tone of the words they speak. Some Japanese speak with a certain "namari" and some with another. "Namari" means a word-tone in a broad sense inclusive of the current intonation and local accents.

What I want to point out is that a tonal transition does not occur on the basis of the vague theory of "long" and "short," or the quantity of sounds, but is it a tonal-dissimilation based on the *quality of sounds*.

For instance, "Tokyo" is not composed of vowels of even pure long vowel sounds but it is a musical form by tonal transition. The accent form of "Tokyo" is usually considered a 'low-high-pattern' and the four syllables contained are in the form of $o/\bar{o} \bar{o} \bar{o}$. It is usually accented like [tōkjō]. The first syllable [to] is pronounced in a low-pitch and the other syllables starting with the next [ō] are in a high-pitch and at the last have a little lower [ō]. In short, the word-tone of "To-" changes as its quality of sound changes. To explain the theory conversely, the vowel sound is not prolonged to make the linguistic-meaning clear but the sound is syllabically dissimilated by transition. The pronunciation of "Osaka" [ōsaka] and "Kyoto" [kjōto] comes under this category.

In conclusion, I should like to point out that the quantity of sounds, which ignores the *quality of sounds* or has nothing to do with the tone, is not realistic, and at the same time I do hope that we have courage enough to realize a new adequate system of notations.

DISCUSSION

Carnochan:

It is useful to distinguish at least two levels of statement here, one for the phenomena, where durations are measurable in centiseconds, and the other for the phonological abstractions. In calling out "Breakfast!" however long one holds on to the first vowel, it cannot be long in quantity. For English, it is the quality that is more important for such distinctions as *i/I*, while different durations for both vowels in different phonetic environments need to be systematised as well, e.g. for heed/heat and hid/hit.

Slis:

The question whether duration is the most important cue in the distinction between so called long and short vowels will be dependent on the language studied. For Dutch for instance it is possible to prove, with synthetic steady state vowels, that a difference in duration only, can cause a change in vowel perception; e.g. a long stimulus with $F_1 = \pm 400$ Hz and $F_2 = \pm 2500$ Hz is perceived as [e]; a short one of the same spectral composition as [ɪ].

Onishi:

a) *On the Duration of Time*, the definition of "Phone" viz. "unit of speech sound" is firstly important. Ordinary, phonetic symbols represent it. And, deliberate or exceptional physiological emitting of sounds has to be excluded.

b) *Long & Short Vowels*, found in for example Indian or Greek languages are the matter of the so-called Phonology, and not of modern scientific phonetics. Again, strictly speaking, there are no pure long vowels.

ANAESTHESIEBEHANDLUNG BEI DYS- UND PARAESTHESIEN IM KEHLKOPFBEREICH

J. PAHN—CH. SEIFERT*

In der phoniatischen Praxis erscheinen nicht selten Patienten mit Stimm-
beschwerden und gleichzeitigen Brennen, Stechen, Kratzen, Schmerz-, Druck-, Ein-
engungs- und Würgeempfindungen im Kehlkopfbereich. Unser besonderes Interesse
galt denjenigen unter diesen Patienten, die keinen organischen Befund zeigten,
die kein „Globusgefühl“ und keine Schluckbeschwerden hatten, die weder neurologisch
noch psychiatrisch besonders auffällig waren und deren Miuempfindungen auch
nach längerer Stimmruhe weiterbestanden. Primäre funktionelle Störungen konnten
als Ursachen ausgeschlossen werden. Im Zweifelsfall entschied der Versuch einer
funktionellen Behandlung, der aber immer erfolglos blieb. Psychogene Reaktionen
und Entwicklungen waren dagegen häufig anzunehmen.

Maximow schlug zur Behandlung solcher und ähnlicher Fälle eine intracutane
Anaesthesia der Head'schen Zonen des Kehlkopfes vor. Wir griffen diese Gedanken
auf und gelangten zu einer Behandlungsmethode, die

1. eine intracutane Anaesthesia über dem Kehlkopf,
2. eine Beseitigung der stimmlichen Funktionsfehler, die sekundär fast immer auf-
traten, und
3. eine Einstellung mit Sedativa umfaßt.

Die intracutane Anaesthesia baut auf der Wirksamkeit viscerocutaner Reflexe auf.
Mittels einer Reizstromtestung werden die überempfindlichen Hautbereiche über
dem Kehlkopf ermittelt und durch eine Kneif- und Kratzprobe nachgeprüft. Nach
Maximow kommen als Zonen erhöhter Hautempfindlichkeit für den Kehlkopf
Felder von 1—2 cm Durchmesser symmetrisch neben den Durchtrittspunkten des
N. laryngeus superior durch die Membrana Hyothyreoidea, symmetrisch neben
der Cart. cricoidea, über der Mittellinie der Cart. thyreoidea und über der Mitte
der Cart. cricoidea in Betracht. Unter diesen Feldern ergeben sich bei jedem Pa-
tienten unterschiedlich einige mit relativ höherer Empfindlichkeit. Für diese fanden
wir als Schwelle der ersten Empfindung des Reizstromes Werte zwischen 0,01 und
2,0 mA. Die Schmerzschwelle, die in Abhängigkeit von einer niedriger oder höher
gelegenen Empfindungsschwelle steigt oder fällt, reichte in diesen Feldern von 0,1

* Aus der HNO-Klinik der Med. Akademie Erfurt (Dir.: Prof. Dr. Kurt Schröder).

bis 8,0 mA. Außerhalb der Hautbezirke mit erhöhter Empfindlichkeit fanden wir Empfindungsschwellen bis zu 3,5 mA und Schmerzschwellen bis zu 14,0 mA. Zur Injektion wurden die Hautfelder mit den relativ niedrigsten Werten ausgewählt.

Wir unterzogen 20 Patienten dieser Behandlung. Davon standen 12 als Dozenten, Lehrer und Erzieher in Berufen mit hoher stimmlicher Belastung. An fünf aufeinanderfolgenden Tagen injizierten wir verteilt auf die ermittelten Hautfelder je 5 cm³ 1% Jecoffin. Bei einigen Patienten wiederholten wir diese Behandlung nach mehreren Monaten. Die Dys- und Paraesthesien ließen meistens schon während der Injektionsbehandlung nach, nur bei wenigen Patienten stellte sich der Erfolg bis zu einem Abstand von mehreren Wochen verzögert ein. Nach stattgefundener Anaesthesiebehandlung erhielten bis auf 5 Patienten alle eine funktionelle Behandlung mit täglichen Sitzungen. Im Endergebnis zählten wir

- 4 Patienten mit vollständiger Heilung,
- 6 Patienten mit fast vollständiger Heilung,
- 8 Patienten mit einer deutlichen Besserung und
- 2 Patienten mit geringem, aber doch noch spürbarem Erfolg.

Wegen der sicher vorhandenen psychogenen Beteiligung ist es ratsam, von vornherein psychotherapeutische Gesichtspunkte zu berücksichtigen, um der Testung, den Injektionen und der funktionellen Übung ein möglichst großes Gewicht zu verleihen, damit die Gelegenheit der vielleicht letztmöglichen Hilfe nicht abgewertet oder verpaßt wird.

DISCUSSION

Böhme:

- Hinweise auf den therapeutischen Effekt des galvanischen Stromes: 1. analgesierend,
- 2. hyperämisierend,
- Anwendung in der Phoniatrie möglich.

Wendler:

Können Sie die angewandte Methode der Empfindlichkeitsprüfung genauer beschreiben?

EINIGE PROBLEME DES PROSARHYTHMUS

ZDENA PALKOVÁ*

Die grundlegende Frage, mit der ich mich gegenwärtig beschäftige, ist die Frage danach, ob in der Prosa eine solche Organisationsart von Lautqualitäten bestehen kann, die man Rhythmus nennen könnte — und wenn sie besteht, ob und wodurch dieser Rhythmus bereits im Autorentext fixiert ist. Im Rahmen eines so umschriebenen Forschungsbereiches müßte eine Reihe von Teilproblemen beantwortet werden. Eines von diesen ist die Frage nach der Deutlichkeit von Lauteinheiten im Tschechischen für den tschechischen Hörer. — Ihr wird dieser Beitrag gewidmet.

Ich gehe dabei von dem Standpunkt aus, daß bei der Prosa, in der der Hörer a priori keinerlei rhythmische Absicht vermutet, nur mit solchen Erscheinungen gerechnet werden kann, die wenigstens in dem Maße deutlich erfaßbar sind, daß man sie testen kann; hierunter verstehe ich: es ist möglich, durch ihre Bestimmung auch vom theoretisch ungeschulten Hörer ihre Identifikation im Material zu erlangen.

Im Tschechischen kommen hierbei zwei Lauteinheiten in Betracht, die in ihrer Ordnung zwei gegenseitig übergeordnete Grade darstellen: der Betonungstakt und der Sprechabschnitt.

Der Betonungstakt ist eine Einheit, die im Prinzip mit dem Versfuß in der klassischen Metrik verglichen werden kann und auf die Existenz des Wortakzents sich stützt (im Tschechischen stets auf der ersten Wortsilbe). Als einziger strittiger Fall bei seiner Bestimmung im Text wird das Vorkommen von einsilbigen Wörtern betrachtet (die manchmal keinen Wortakzent zu haben brauchen). Als Einheit, vermittels derer es möglich ist, den rhythmischen Verlauf der Mitteilung zu charakterisieren, ist er allgemein anerkannt.

Der Sprechabschnitt — die dem Takt unmittelbar übergeordnete Einheit — basiert auf keiner solchen grammatisch charakterisierbaren Erscheinung. Als Lauteinheit wurde er vor allem in der Arbeit von F. Daneš *Intonace a věta ve spisovné češtině* (Intonation und Satz in der tschechischen Schriftsprache), Praha 1957, formuliert. Es handelt sich um eine Taktgruppe, die als Lauteinheit (vom rhythmischen Gesichtspunkt aus betrachtet) durch die Anwesenheit eines sog. Intonationszentrums bestimmt werden soll, eines Höhepunktes, zu deren Realisation Intensität,

* Phonetisches Institut der Karlsuniversität, Prag.

melodische Kadenz und eventuell die Möglichkeit einer darauffolgenden Pause beiträgt. Apriorisch läßt sich die Realisation dieser Einheit als eine bloß potentielle verstehen (man kann z. B. die Wortakzente in einem bestimmten Text so stark realisieren, daß jeder Takt selbständig bleibt). Es handelt sich demgemäß um eine Einheit, deren Wesen dem Laien schlechter deutlich gemacht werden kann als der Takt. Den Grad ihrer Klarheit für den Hörer habe ich vor allem überprüfen wollen.

Zwölf verschiedene Texte von 5 tschechischen Prosaikern in der Interpretation von 9 verschiedenen Sprechern mit etwa 350 Einheiten beider Typs bewerteten in 2 Testreihen stets 15 tschechische Hörer; hierbei dürften als grundlegend die 8 Antworten von den Hörern bewertet werden, die beide Testreihen absolviert haben. Alle Hörer waren theoretisch unbefangene und man konnte bei ihnen eine geläufige Lese-Erfahrung voraussetzen.

Ich möchte über zwei Ergebnisse sprechen, die aus dem derzeitigen Material hervorgingen:

I. Es scheint, daß die höhere Einheit — der Sprechabschnitt — obgleich er vage eingegrenzt ist — deutlicher spürbar ist, als der Betonungstakt.

Die Bestimmung der Betonungstakte fiel unsicherer aus als man hatte annehmen dürfen. Bei den einsilbigen Wörtern blieb 1/3 der Fälle unentschieden (es sind nicht einmal Übereinstimmungen von $p = 0,6$ bei ihrer Bestimmung erreicht). (Die Lage bei einsilbigen Wörtern wurde noch gesondert getestet, zur Verfügung standen die Ergebnisse von 25 Hörern). Die Unsicherheit betraf jedoch auch mehrsilbige Takte, am häufigsten zweisilbige. Im Ganzen:

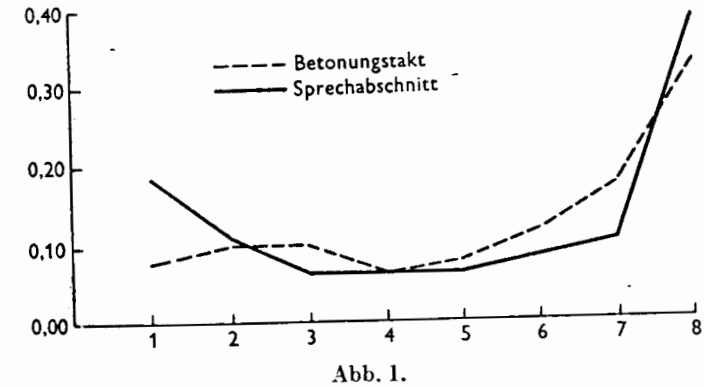
Bei der Forderung von 3/4 Übereinstimmung bei den Hörern blieb die Segmentierung in Takte ungenau bei etwa 1/3 des Textes, die Segmentierung in Abschnitte bei 1/4 des Textes (vgl. die Anlage; weitere Zahlangaben sind vorhanden). Das Verhältnis der übereinstimmenden Bewertung im Vergleich zu der überhaupt durchgeführten Textbewertung war dabei bei beiden Einheiten nahezu gleich — 1 : 3 bis 1 : 4. Die Bewertung der Abschnitte schöpft allerdings aus einem wesentlich größeren Komplex theoretisch vorstellbarer Möglichkeiten: der Anteil übereinstimmender Bewertungen umfaßte bei unserem Testmaterial im Falle der Takte die Auswahl der Hälfte aller Möglichkeiten, im Falle der Abschnitte die Auswahl eines Fünftels. (Kennzeichnenderweise erhöht der Anteil aller vorkommenden Bewertungen bei den Takten um 14% den Anteil der theoretisch annehmbaren Möglichkeiten.) — Auch individuelle Abweichungen in der Bewertung der Takte sind höher als in der Bewertung der Abschnitte. Die Ergebnisse der Hörer, deren Bewertung im Takt-Test deutlich vom Durchschnitt abstach, glichen sich im Test der Abschnitte dem Durchschnitt an.

Es scheint demnach vorerst, daß der Sprechabschnitt eine genügend deutlich erfaßbare Einheit für den tschechischen Hörer darstellt, so daß man ihn als den eventuellen Träger des Prosa-Rhythmus im Tschechischen in Betracht ziehen

kann. Vom Gesichtspunkt der Deutlichkeit ist er eine markantere Einheit als der elementare Betonungs-Takt.

II. Bei der Beurteilung der Konstituierung einer Lauteinheit, von der wir voraussetzen, daß sie als rhythmische Einheit dienen kann, ist nicht nur ihr Höhepunkt zu berücksichtigen, sondern auch Charakteristika, die ihre Grenze angeben.

Das Material wurde zur Bewertung in den Testen auf zweierlei Art vorgelegt. Die Ergebnisse zeigten die Abhängigkeit der Antworten hiervon. Anschaulich waren die Ergebnisse bei dem Test der Gruppe monosyllabischer Wörter, die allgemein als unbetont betrachtet werden.



Im Test A bestimmten die Hörer primär die Betontheit oder Unbetontheit beim Einsilber, sekundär seine Zugehörigkeit zu einem der Nachbar-Takte. Im Test B teilten sie umgekehrt die zusammenhängende Reihe der Mitteilung in Takte auf und bestimmten erst in diesen den Akzent-Gipfel. Die Ergebnisse in der Bestimmung der einzelnen Wörter unterscheiden sich grundlegend. Mehr als 30% der Bewertungen verschoben sich und nahezu die Hälfte der getesteten Wörter.

Die Abweichung betraf dabei vor allem die Bestimmung des Höhepunktes und nur in geringerem Maße die Bestimmung der Zwischentakt-Grenze. Im Test B erhöhte sich die Wahrscheinlichkeit, daß der Einsilber keinen Akzent trägt (das heißt im Einklang mit der theoretischen Annahme), und zwar in gleichem Maße für beide Möglichkeiten, ihn in die Nachbartakte einzugliedern. Im Test B zeigte sich auch eine um 20% höhere Übereinstimmung der Hörer-Urteile. — Der Test der Bestimmung der Sprechabschnitte wurde einem ähnlichen Versuch unterzogen. Die Bedeutung der Grenze der Einheit trat dabei noch deutlicher zutage. Das Abschnitts-Zentrum — d. h. den Akzent-Höhepunkt des Abschnittes — zu testen, ist bisher nicht gelungen.

Es scheint demnach vorerst, daß der Hörer die Grenze von Lauteinheiten im Prosa-Text sich deutlicher vergegenwärtigt, oder anders, daß er den Gipfel einer rhythmischen Einheit oft erst in ihrem Rahmen identifiziert, nicht die Einheit an ihrem Gipfel. Es ergibt sich die Frage nach weiteren Lautqualitäten, die diese Einheit als Einheit einschränken.

Die Untersuchung schreitet in allen Teilen fort.

Tab. 1. Übereinstimmung in der Bewertung der Einheits-Grenze (Wahrscheinlichkeitswerte)

Zahl der übereinstimmenden Beurteilungen Einheit	1	2	3	4	5	6	7	8
T	0,07	0,09	0,09	0,06	0,07	0,11	0,17	0,33
A	0,18	0,10	0,06	0,06	0,06	0,08	0,10	0,37

T — Betonungstakt

A — Sprechabschnitt

2. [1] in graphischer Darstellung

ZUR PHYLOGENETISCHEN ENTWICKLUNG DES KEHLKOPFSKELETTES VON DEN AMPHIBIEN BIS ZUM MENSCHEN

K. PAULSEN*

Die phylogenetische Entwicklung des Larynxskelettes ist zuletzt ausführlich 1936 von Göppert dargestellt worden. Schon lange vor dieser Zeit waren die wesentlichen Entwicklungsstufen des Kehlkopfes bekannt. Man hatte eine fortschreitende Differenzierung von den Amphibien bis zum Menschen festgestellt und unterschied das primäre Larynxskelett vom sekundären, dem Kehlkopfgerüst der Säugetiere und des Menschen. Bis heute hat sich an dieser Darstellung der Phylogenese nichts geändert.

Die letztlich die Stimmbildung bestimmenden Einrichtungen wie die Stellung der Larynxknorpel zueinander, besonders aber die Stellung der Aryknorpel und die Lage der Stimmbänder sind niemals gebührend berücksichtigt worden. Für den an der Stimmbildung Interessierten scheint aber eine derartige funktionell-anatomische Betrachtung besonders wichtig. Die Leistung eines Stimmorganes wird wesentlich von der Lage der Stellknorpel und der Zuordnung der Stimmbänder im Kehlkopfgewebe bestimmt. Da trotz zahlreicher artgebundener Variationen jede Wirbeltiergruppe einer bestimmten Entwicklungsstufe entspricht, lassen sich durch den als gemeinsam erkannten Bauplan des Larynxgerüsts wichtige Rückschlüsse auf die Stimmleistung ziehen.

Im folgenden möchte ich Ihnen die Phylogenese des Larynxskelettes in der von mir genannten Betrachtungsweise entwickeln. Die große Gruppe der Fische scheidet aus. Wie weit man die Dipnoer, die Lungenfische, berücksichtigen sollte, läßt sich schwer entscheiden. Ich beginne daher mit den niedersten Landwirbeltieren, den Amphibien.

Das Urskelett des Kehlkopfes treffen wir unter den niederen Amphibien, den Urodelen, an. Beiderseits der Atemöffnung liegt ein Knorpel, die sog. Cart. lateralis, aus der später die drei Anteile des primären Larynxskelettes, die Arytaenoide, das Cricoid und die Trachealknorpel hervorgehen. Obwohl Stimmbänder fehlen, erzeugen diese Tiere schon Laute, die allerdings recht leise und in ihrer Modulationsfähigkeit stark eingeschränkt sind. Als Generatoren solcher Stimmen dürften die Schleimhautwülste des Aditus laryngis fungieren.

* HNO-Klinik der Universität Kiel.

Richtige Stimmbänder besitzen erst die Anuren, die höheren Amphibien. Bei diesen Tieren hat eine erste Teilung des Kehlkopfgerüsts eingesetzt: selbständige Arytaenoide sitzen einem unpaaren Knorpel, der Cartilago trachealis, breitbasig auf. Die Arytaenoide sind dreieckförmig gestaltet und schließen durch ihre konvexe Form den Larynx nach außen hin praktisch ab. Die höheren Amphibien sind außerordentlich stimmbegabt. Ihre Stimmen sind im allgemeinen recht niederfrequent und nur wenig modulierbar, aber sehr kräftig und tönend. Diese Eigenschaften sind offensichtlich an Stimmbänder — oder jedenfalls leicht schwingungsfähige Strukturen — gebunden. Charakteristisch ist die Anheftung der Stimmlippen im Larynx: Sie sind immer entlang der Basis der Stellknorpel ausgespannt, also des Knorpels, der sich im Kehlkopf der Amphibien allein ausreichend verformen läßt. Da sich durch die Anheftung an einem einzigen Knorpel nur die Spannung, kaum aber ihre Form wesentlich verändern läßt, sind der Lautbildung recht enge Grenzen gesetzt.

In der nächsthöheren Gruppe der Wirbeltiere, den Reptilien, ist die Dreiteilung des primären Larynxskelettes in Arytaenoide, Cricoid und Trachealknorpel endgültig vollzogen. Das Cricoid ist auf seiner ventralen Seite höher gestaltet, das Arytaenoid nach dorsal versetzt und um seine Achse gekippt. Die Gruppierung der Muskulatur ist neu geordnet. Obwohl die Sphinkterfunktion noch deutlich erhalten ist, ist die Anordnung differenzierter als bei den Amphibien. Stimmbänder sind nur in wenigen Familien ausgebildet. Sie sind regelmäßig ausgespannt zwischen Cricoid und Stellknorpel. Die Stimmbänder sind wie bei den Amphibien lediglich membranös gestaltet, jedoch wesentlich kräftiger.

Dadurch, daß die Arytaenoide nur für einen Teil des Stimmbandes als Fixationspunkt dienen, scheint die Leistungsfähigkeit solcher Stimmorgane für die Stimmbildung größer als bei den Amphibien. Die Spannung der Stimmlippen wird sich durch die Verschiebung der Stellknorpel im Cricoid-Arytaenoid-Gelenk feiner abstufen lassen, als die Spannung bei den nur an den Arytaenoiden angehefteten membranösen Stimmlippen der Amphibien. Leider liegen über die Stimmäußerungen der Reptilien bisher nur spärliche Befunde vor, so daß sich über die Leistungsfähigkeit ihrer Stimmorgane etwa im Vergleich zu denen der Amphibien kaum urteilen läßt.

In der Phylogenese der Reptilien werden zwei Gruppen unterschieden, die für die weitere Entwicklung der Wirbeltiere wegweisend sind: die Theropsiden als Vorfahren der Säugetiere und die Sauropsiden als Vorfahren der Vögel. Betrachten wir zunächst das Larynxskelett der Vögel. Ein wesentlicher Unterschied gegenüber den Reptilien fällt nicht auf. Im ganzen sind die Kehlkopfanteile zarter gebaut, das Arytaenoid scheint noch weiter um seine Achse gedreht und nach dorsal versetzt, das Cricoid noch höher. Alles in allem sind es nur geringfügige Veränderungen. Das besondere Merkmal am Larynx der Vögel sind die fehlenden Stimmbänder. Damit ist in dieser besonders stimmbegabten Wirbeltierklasse der Kehlkopf als das Organ der Lauterzeugung verlassen. Statt dessen wird die Stimme in der tiefergelegenen und viel leistungsfähigeren Syrinx gebildet.

Die Säugetiere dagegen haben den Kehlkopf als Organ der Stimmbildung beibehalten. Gegenüber dem bisher genannten ist im Aufbau jedoch ein deutlicher Fortschritt sichtbar. Neben dem Erwerb eines vierten Knorpels, der Cartilago thyroidea, also des Schildknorpels, ist — und das erscheint mir besonders wichtig — das Arytaenoid im Verhältnis zu den übrigen Anteilen klein, noch stärker um seine Achse gedreht und ganz auf die dorsale Larynxfläche gerückt. In seiner kugelgelenkartigen Verbindung mit der Ringplatte des Cricoids ist es in jeder Richtung beweglich. Diese Eigenschaften erlauben eine schnelle und fein abstufbare Einstellung der immer zwischen Schild- und Aryknorpel ausgespannten Stimmbänder, die punktförmig am Processus vocalis des Stellknorpels angeheftet sind. Für die Stimmbildung muß dieser Mechanismus wesentlich leistungsfähiger als der der Amphibien und Reptilien sein.

Zum Schluß sehen Sie Phylogenese des Larynxskelettes noch einmal schematisch in einer Übersicht zusammengefaßt. Das Wichtigste und auch in der Entwicklung der Stimmbildung Bedeutsamste ist die Wanderung des Aryknorpels nach dorsal mitsamt seiner Drehung um die eigene Achse. Als beweglicher Fixationspunkt bestimmt überwiegend der Stellknorpel Stellung und Spannung der Stimmbänder. So läßt seine Lage zu den übrigen Kehlkopfanteilen einen gewissen Rückschluß auf die Leistungsfähigkeit einer Tierstimme zu.

Selbstverständlich wird die Leistungsfähigkeit eines Stimmorgans nicht allein durch die eben dargestellten anatomischen Gegebenheiten bestimmt.

AUTOREFERAT ZUM FILM „DIE STIMMBILDUNG IN DER AUFSTEIGENDEN WIRBELTIERREIHE“

K. PAULSEN

Der Film wird eingeleitet mit einem Trickschema über die Entwicklung der Wirbeltiere von den Fischen bis zum Menschen. Ein weiterer Trick zeigt den Mechanismus der Stimmbildung, die im Prinzip der einer Zungenpfeife gleicht. Der Zunge der Pfeife entsprechen die Stimmbänder, dem Windraum die Lunge, dem Ansatzrohr der Rachen, Mund- und Nasenraum (dabei Überblendung des Bildes der Zungenpfeife mit einem auf das Grundprinzip beschränkten Bild eines Phonationsorganes der Landwirbeltiere).

Im ersten Hauptteil des Filmes wird die Phylogenese des Larynxskelettes von den Amphibien bis zum Menschen mit Trickzeichnungen dargestellt. Die Urform des Kehlkopfskelettes bei den niederen Amphibien besteht aus einem paarigen Knorpel, der *Cartilago lateralis*. Aus dieser Urform des primären Larynxskelettes entwickelt sich bei den höheren Amphibien das einteilige Cricoid mit den breiten selbständigen dreieckförmigen Arytaenoiden, an deren Innenflächen Stimmfalten ausgespannt sind. In der weiteren Entwicklung richtet sich das Arytaenoid auf. Die Dreiteilung des primären Larynxskelettes ist bei den Reptilien vollzogen. In dieser Tierklasse liegen die Stimmbänder zwischen dem Arytaenoid und dem Cricoid. Der Kehlkopf der Vögel ist wenig differenzierter als der der Reptilien. Stimmbänder fehlen hier immer. Als Phonationsorgan dient die Syrinx mit schwingenden Paukenmembranen. Bei den Säugetieren und beim Menschen hat sich durch Hinzutreten eines vierten Kehlkopfknorpels, des Thyreoids, das sekundäre Larynxskelett ausgebildet. Auffällig ist hier die starke Dorsalverlagerung des Arytaenoids und die starke Kippung um mehr als 90° um die Achse. Stimmbänder sind ausgespannt zwischen Arytaenoid und Thyreoid.

Der zweite Hauptteil, der den Mechanismus der Stimmbildung bei den Amphibien am Beispiel des Frosches, bei den Reptilien am Beispiel des Geckos, bei den Vögeln am Beispiel der Gans und bei den Säugetieren am Beispiel des Menschen zeigt, gliedert sich in vier Unterabteilungen. Jeweils wird zunächst die genannte Tierart in ihrem Biotop gezeigt. In der Folge wird auf das im ersten Hauptteil dargestellte Larynxskelett zurückgeblendet; die dritte Unterabteilung zeigt schematisch die Stimmbildung im Querschnitt des jeweiligen Stimmorganes in der Zeitdehnung. Die vierte Unterabteilung bringt die Originalaufnahmen der Hochgeschwindigkeitsfilme in Zeitdehnung.

Im einzelnen wird gezeigt: 1. quakende Frösche im Biotop. Rückblendung auf das Kehlkopfschema und anschließend Trickfilm über die schwingenden Stimmbänder des Frosches anhand eines Schemas. Dabei sieht man, daß sich die am Glottisrand wulstig verdickten, sonst zarten, membranösen Stimmbänder breit aneinanderlegen, daß der Luftstrom zunächst die seitlichen Stimmbandanteile segelartig aufbläht und schließlich die Glottis sprengt. Vermöge ihrer Elastizität schwingen die Stimmbänder schnell wieder in ihre Ausgangslage zurück. Die Stimmbänder schwingen senkrecht zum Luftstrom. Im vierten Teil wird eine Zeitdehneraufnahme von einem Präparat der schwingenden Stimmlippen des Frosches bei einer Bildfrequenz von 8 150 Bildern in der Sekunde gezeigt. Man blickt dabei durch den geöffneten Aditus auf die geschlossenen Stimmbänder des Frosches. Die Stimmfrequenz beträgt etwa 20 Hz. Die Verschußphase ist außerordentlich lang, die Schwingungsphase dagegen sehr kurz. Der Frosch ändert die Tonhöhe seiner Stimme nur durch Verkürzen oder Verlängern der Verschußphase. Die Dauer der Schwingungsphase bleibt immer konstant.

2. Der Tokee-Gecko im Biotop. Der knarrende Schrei wird nur bei weitgeöffnetem Maul ausgestoßen. Rückblendung auf das Kehlkopfskelett-Schema der Reptilien. Im dritten Teil Trickfilm über die Stimmbandschwingungen des Geckos anhand eines Schemas in Zeitdehneraufnahme. Während der Phonation steht die untere Fläche der Stimmbänder senkrecht zum Luftstrom. Gering gewölbte Randgebiete werden zunächst mit Luft gefüllt, bis der steigende Druck die Glottis aufreißt. In der Schwingungsphase bewegen sich auch die Kehlkopfwandungen stark mit. Die Stimmbänder schwingen vorwiegend senkrecht zum Luftstrom. Die Originalaufnahmen von den schwingenden Stimmbändern eines angeblasenen Larynxpräparates des Geckos zeigen bei einer Zeitdehnung von 8 000 Bildern in der Sekunde den Schwingungsablauf bei einer Stimmfrequenz von 300 Hz. Die gezeigten Aufnahmen entsprechen nicht ganz dem natürlichen Phonationsablauf, da sich die muskulösen Stimmbänder am Larynxpräparat nur annähernd so spannen ließen wie beim lebenden Tier. Die Verschußphase erscheint kurz. Auch bei den Reptilien ist die Schwingungsphase konstant. Auch der Gecko ändert seine Stimmfrequenz nur durch Verkürzen oder Verlängern der Verschußphase.

3. Schnatternde Gänse im Biotop. Im weiteren wird auf das Larynxskelett der Vögel rückgeblendet und ein Trickschema von den schwingenden Membranen der Syrinx gezeigt. Die schwingenden Paukenmembranen der angeblasenen Gänsesyrix werden bei einer Bildgeschwindigkeit von 5 500 Bildern in der Sekunde als Hochgeschwindigkeitsaufnahmen in Zeitdehnung aus verschiedenen Blickrichtungen vorgeführt. Die Schwingungsfrequenz beträgt etwa 400 Hz. Die Hochgeschwindigkeitsfilme werden in folgender Reihenfolge gezeigt: schwingende Paukenmembranen mit erhaltenem Bronchidesmus von hinten außen, schwingende Paukenmembranen mit gelöstem Bronchidesmus von hinten außen, schwingende linke äußere Paukenmembran und schwingende innere Paukenmembran von oben nach Abtragung der Trachea. Die Paukenmembranen schwingen rhythmisch hin und her, wobei äußere

und innere Membranen den gleichen Takt aufweisen. Der Hauptbronchus führt gegen die Trachea eine Art Kipp-Kreis-Bewegung aus, die immer einem bestimmten Bewegungsabschnitt der schwingenden Membran zugeordnet ist. Innere und äußere Membran berühren sich während der Schwingung nicht. Bei der Phonation bleibt also immer ein Spalt zwischen beiden Paukenmembranen bestehen.

4. Rückblendung auf das Larynxskelett der Säuger und des Menschen. Im Trickfilm werden dann anhand eines Schemas die Schwingungen der Stimmbänder in Zeitdehnung vorgeführt. Dabei wird lediglich die Veränderung an der Schleimhautoberfläche gezeigt, während die Schwingung des gesamten Stimmbandes nur anhand einer in die Stimmbänder eingezeichneten ellipsenförmigen Figur wiedergegeben wird. Die Stimmbänder schwingen nicht als Ganzes wie die Stimmbänder des Frosches und des Geckos. Vielmehr werden ihre einzelnen Teile auf einer elliptischen Bahn nacheinander von der Impulswelle verlagert. Die Originalaufnahmen von den schwingenden Stimmbändern eines phonierenden Mannes werden in Zeitdehnung — bei etwa 3 000 Bildern in der Sekunde — bei einer Stimmfrequenz von 140 Hz gezeigt. Der Verlauf der Impulswelle läßt sich gut verfolgen. Im Gegensatz zu den Amphibien und Reptilien wird bei den Säugetieren wie beim Menschen die Tonhöhe durch Verlängern oder durch Verkürzen der Schwingungsphase bestimmt. Auch während der Verschußphase ist das Stimmband in Bewegung.

Zum Schluß des Filmes wird noch einmal in kurzer Folge der Schwingungsablauf beim Frosch, beim Gecko, bei der Gans und beim Menschen gezeigt.

ZUR LAUTWIEDERGABE IN LEHNWÖRTERN DES ALTHOCHDEUTSCHEN

H. PENZL*

Der Vergleich von Lehnwörtern in ahd. Dialekten mit ihren lateinisch-romanischen Ursprungsformen sollte für Nehmersprachen wie Gebersprachen phonetisches Beobachtungsmaterial ergeben. Wichtig ist hiebei vor allem, daß man nicht außer acht läßt, daß es bei der Lautwiedergabe in der Nehmersprache zweierlei Arten von Lautsubstitution gibt.¹ Die eine ist Lautnachahmung, in der der Laut der Gebersprache durch den phonetisch nahestehendsten Laut wiedergegeben wird und die andere Lautersatz, bei dem nicht nur phonetische, sondern vor allem morphophonemische, phonotaktische, morphemische Faktoren mitspielen. Bei teilweiser schriftlicher Entlehnung, neben oder statt der mündlichen, zeigt die Wiedergabe der Nehmersprachen oft Züge einer alphabetischen Umsetzung oder einer Transkription.

Diagraphischer Vergleich der Wiedergabe lateinischer Formen und Laute in ahd. Texten verschiedener Zeit und Dialekte hat gezeigt, daß Unterschiede der Wiedergabe oft auch Lautentwicklung der Gebersprache sowie mehrfache Entlehnung durch die Nehmersprache widerspiegeln. Wie stellt sich nun im einzelnen das Lautsystem der Nehmersprache (oder des Nehmerdialektes) zur Wiedergabe lateinisch-romanischer Lautwerte? Wir wollen hier die Lehnwörter in Otfrids südrheinfränkischer Evengelienharmonie aus der zweiten Hälfte des 9. Jhds. untersuchen.

Otfrid erwähnt ausdrücklich in seinem Schreiben an Erzbischof Liutbert, daß er den Wert der lat. Vokalzeichen *a e i o u* in seinem Fränkischen beibehalte. Sein bewußter Gebrauch von *y* im Präfix *yr*, den Digraphien *ya ye* (*sýazo*, *blýent*) findet durch seine nur orthographisch bedingte Verwendung in Fremdwörtern keine weitere Aufklärung: *elemosyna*, *myrrun*, *paradys* neben *paradis*, usw.² Der Vergleich seiner Vokale mit den vermutlichen Lautwerten in den Gebeformen zeigt überwiegend Lautnachahmung: *altari*, *fundament*, *gigant*, *karitas*, *corona*, *lilia*, *mandat*, *metar*, *musica*, *nardon*, *natura*, *nona*, *organa*, *prosum*, *regula*, *rosa*, *sextari*, *sillaba*. Daneben gibt es Anzeichen von ehemaligem Lautersatz, z. B. *i u* statt *e o* vor Nasal oder *i* der Nebensilbe: *gimma* 'gemma', *zins* 'census' neben *sens* 'sensus'; *sichur*

* University of California, Berkeley.

¹ „Namen und althochdeutsche Lautverschiebung“, *Proceedings of the Eighth International Congress of Onomastic Sciences* (Haag 1966), S. 384 ff.

² Joh. Kelle, *Die Formen- und Lautlehre der Sprache Otfrids* (Regensburg 1869), S. 445 f.

'securus', *munizara* 'Wechsler (Pl.)'. Das Umlautsphonem /e/ vor *i* wie in *engil* 'angelus', *esil* 'asinus', *kelih* 'calix', *kestiga* 'castigo', *ketina* 'catena' bedeutet keine Übernahme vor der Umlautperiode, sondern ist einfach Lautersatz des palatalen *a* durch das phonotaktisch übliche /e/ vor /i/. Dieser konnte jederzeit, auch nach Vollendung des Umlautes eintreten, läßt also keinerlei Schlüsse auf die Chronologie der Übernahme oder gar der ahd. Lautentwicklung (gegenüber Isidors *angil*?) zu.

Bei der Wiedergabe der lat. Sonorlaute gibt es dieselben Zeichen *l r m n* wie im Lateinischen: *regula*, *castelle* (Dat.), *corona*, *myrrun* usw. Ein dissimilatorischer Wechsel von *l* und *r* deutet auf phonetische Verwandtschaft der beiden Liquiden, also im Dialekte Otrfrids auf apikales [r], nicht uvulares [R]: *martolotun* 'sie marterten' (bairische Freising Hs.: *martorotun*) neben *martyra*; *murmulo* (3. Sing. Konj.) 'murmurare', *murmulunga*.

Im Obstruentensystem Otrfrids stehen zwei Reihen von Spiranten in Opposition, ebenso zwei Reihen von Verschlusslauten, die sich durch Fortis-, bzw. Lenisartikulation unterscheiden. Bei den Spiranten ist die Opposition auf den Inlaut beschränkt, im Anlaut steht Lenis, im Auslaut Fortis. Dazu stimmt die Verteilung von *v ff s h hh* in den Lehnwörtern.

Unter den Verschlusslauten stehen labiale Lenis und Fortis im Anlaut in Opposition: *pad* 'Pfad', *bad* 'Bad'. Lehnwörter mit lat. *p* zeigen *p* und *b*: *spenton*, *palinza*, *palma*, *pina*, *pinon*, *helliporta*, *prosun*, *pruanta*, *purpurin*, *scorpion*; *beh*, *brediga*, *bredigari*, *bredigon*, in Hs. F: *hellibina*. Beides ist wohl Lautnachahmung aus verschiedenen Gebeformen und verschiedener Tradition: romanische stimmlose Lenis, bzw. nichtaspirierte Halbfortis wurde mit /p/ und /b/ nachgeahmt. Im Inlaut ergibt *p* in *purpurin*, *scorpion* gegen (b) in *uuurbun* 'warben', *korbi* 'Körbe' eine sonst nicht belegte Opposition, da dort die Affrikata *ph* (*ɸph*) dem *p* des Anlautes entspricht: *harpha*, *ophar*, Gen. *oppheres*, *kuphar*, *krippha*.

Die dentalen Verschlusslaute (t) und (d) stehen im In- und Auslaut, nicht im Anlaut in Opposition: *noto* und *odo*, *bat* und *bad*, auch *nardon* und *martyra*. Wir interpretieren daher die direkte Variation zwischen *dunichun* und *tunichun*, *dreso* und *treso*, *dihon* und umgekehrtem *tihonnes* als Schreibungen für die stimmlose Lenis /d/, also als Lautersatz, nicht etwa als Lautnachahmung mit dem /t/ des In- und Auslautes oder gar als Lautübernahme mit dialektfremder Halbfortis.

Die Fortisspirans scheint auf vorahd. Lehnwörter beschränkt: *straza* 'Strasse'. Die Affrikata /tz/ ist meist Lautnachahmung einer romanischen Affrikata, nicht etwa Lautentwicklung aus vorahd. **t*: *kruzi* (*cruci*, *kruci*) 'crucem', *zins* (Hs. F: *cins*), *lekza* 'lectio', *puzzi* 'puteus', *palinza*, *porzicha* 'porticus' (Pl.) mit *t* vor *i*, aber *forta* mit *t* vor Velarvokal.

Die velaren Verschlusslaute /g/ und /k/ stehen in Anlaut und Inlaut in Opposition: vgl. *gimma* und *kamaru* (Dat.), *prediga* und *musica*. Otrfrids Gebrauch von *k* statt *c* in Worten wie *keisor*, *karkari*, *lekza*, *krist*, sogar in *karitas* neben *caritas* spricht deutlich für Lautsubstitution durch das fränkische stark aspirierte *k*, da er in seinem

Schreiben an Liutbert das Zeichen „k“ ob *fautium sonoritatem* („wegen des Schalles in der Kehle“) als notwendig erklärt.

Unsere Untersuchung der Form der lateinisch-romanischen Lehnwörter in Otrfrids Evangelienharmonie hat ergeben, daß zwar verschiedene Lehnwortschichten durch ihre Wiedergabe deutlich erkennbar sind, aber wie von jeher Lautnachahmung und Lautersatz innerhalb des gegebenen Lautsystems durchgeführt wurden. Für Lautübernahme statt Lautsubstitution, ähnlich der bei den französischen Nasallauten in den französischen Lehnwörtern des Neuhochdeutschen, finden wir hier wie überhaupt in der ahd. Periode keinerlei Anzeichen.

DISCUSSION

Blandár:

Die Adaptierung der lautlichen Form ist nur eine Seite des Prozesses der Eingliederung des Lehnwortes in die neue Wortstruktur. Eine besondere Problematik aus diesem Fragekreis in dem bilinguistischen Sprachmilieu. z.B. die ältesten slowakischen Eigennamen in lat., ung., deutschen Sprachdenkmälern. Wenn der Schreiber ein Ungar war, die slowakischen Wörter wurden auf Grund des Lautsystems wiedergegeben. Solche Belege bieten oft Möglichkeiten für die Chronologie der Wortentlehnung und somit für die Siedlungsgeschichte.

Wiesinger:

Das Problem der absoluten und relativen Chronologie von Lautwandlungen ist mit Hilfe nur schriftlich überlieferter Wörter oftmals schwer zu lösen. Bessere Aufschlüsse gewähren die von den Zeiten der ersten Berührung bis in die Gegenwart reichenden, wechselseitigen Entlehnungen im Sprachgrenzgebiet, deren mundartliche Formen die Lautgeschichte widerspiegeln. Dort vermögen auch Ortsnamen mit ursprünglich gleichem, heute aber unterschiedlichem Lautstand die verschiedenen Entlehnzeiten und damit wieder die Lautgeschichte der Entlehnsprache aufzuzeigen. So finden sich im südbairischen Tirol für lateinisch-romanisches *v*, das im Romanischen bis in die Gegenwart gilt, deutsch-mundartlich *t*, *w*, *b*, z.B. für Lat.rom. Venosta—Vintschgau [t], Vipidenum—Wipptal [w], Verona—Bern(e) [b].

REDUCING A PERSEVERATION OF IMMATURE TIMING BETWEEN AN ACOUSTIC SPEECH SIGNAL AND ITS PERCEPTION

GRACE PETITCLERC

An inscription on an ancient Oriental temple says: He who would straighten the end of a process must commence by making the beginning correct.

Consequently, in attempting to reduce the perseveration of immature timing between an acoustic speech signal and its perception it is necessary to re-view the process of speech development when a time lapse between the signal and its perception was normal. This occurs, as you know, in very young children. The acoustic signal enters receptive hearing areas, the attention is caught and held upon the vibrating signal until a translation of the signal into meaningful images takes place. Thus, the signal is perceived.

The time consumed by this process of translation when the child is still very inexperienced gradually diminishes through constant practice until the whole metamorphosis from symbolic signal to image, or perception, becomes instantaneous. That is, unless some physical or psychological barrier prevents maturation. Then, to reduce, or circumvent, the barrier, it is necessary to commence at the beginning and direct the practice of the complete learning process, step by step, into mature instantaneous perception.

Research has disclosed that the first receptors of sound in a young child are the cells of his body. Body-hearing receives and registers vibrations at a level above and below the audible threshold. This hearing is now being accepted as the basis on which future ear-hearing is built. It is also the basis for memory and the focused ability to listen well.

In early childhood there is direct stimulation to the body from clothing, bathing, pin pricks and such, therefore, to supply these same experiences to older children is often impossible. Yet, similar sensory experiences can be achieved with other materials as: rub sandpaper over the hands—the palms, the backs, between the fingers—then over the arms, neck, cheeks, tip of the nose and the lips. Close the eyes during these exercises to aid concentration. And follow one exercise with a contrast of experience, as, after sandpaper rub the same areas with cotton. Or walk barefoot over hot rocky surfaces, then splash in cool mud.

After each stimulating experience translate the memory of it onto a piece of paper with colored chalk. The result may be nothing more than a scribble or a few lines

but to *translate* the sensation into any kind of image commences the process of perceiving the speech signal in translated images, or meaningful concepts.

As progress is made, call attention to the right and left side of the body and determine which side received the stimulation the clearest, or which side distorted it. Encourage the expression of this difference in the translation on the paper. Begin encouraging any element of choice of discrimination, for each added attention-holder aids greater concentration, therefore more accurate memory. Concentrating on the likenesses and differences of the right and left side also develops balance and spatial concepts, as well as hand dominance and body image.

Explore thoroughly all possible facets of body stimulation with inaudible vibrations, then advance into audible sound. Use a vibrator and cover the same areas of the body stimulated before, or have the child lie prone on the floor, make a loud hum in his throat while the therapist pounds over his body with clenched fists. Lying prone over a piano while someone plays very loudly, or over an organ. Even a radio turned up very loud is good.

Next add taste and smell, again increasing the length of concentration, the skill of imagery in meaningful concepts and the speed of producing a result from stimulation to an image. All, of which, is bringing the time element under control to be readily resolved entirely.

While chewing, crunching, gulping, drinking and relishing all sorts of contrasting flavors and textures in foods, listen and feel intently. Determine the volume of food inside the mouth by the feel of it. Place the hands on either jaw and add the consciousness of the sound, the feel, the taste and smell and the muscle movement. Translated on paper, embryonic forms in the expression should be emerging.

Lastly introduce rhythm and visual perception. The rhythmic pattern of speech is the last for the young child, or the perseverator, to master. And rhythm is necessary to meaning.

Begin with simple hand clapping rhythms, advance to complex ones and add music, encouraging in listening the separation of the music melody and the beat. Add any element that inspires intent and discriminatory listening.

The translation can now be varied, using different mediums on paper or other interesting materials. Even colored lights are an excellent variation. Try movement of the body with music, the reproduction of pitch and volume by the voice to accompany the body and the music. This accomplished, augment the process toward complete and instantaneous timing with visual illustrations that tell a story applicable to the music or the body movement or the rhythm. Inspire the child to make the application of the visual segment to the whole.

Thus all of the senses are working together as a team. When they are brought together into an intense unified focus the perseveration of immature timing between an acoustic speech signal and its perception has disappeared.

SUR LA PROBABILITÉ DE DÉTECTION („DETECTABILITY“), DE DIFFÉRENTS MESSAGES VOCAUX

O. PIGNATARO*—G. P. TEATINI*

La méthode d'étude que l'on emploie généralement pour l'application de la théorie de la décision à la réception des messages vocaux est basée sur un procès à deux stades: dans le premier stade on présente au sujet un signal (généralement un mot) qu'il doit identifier; dans le second stade le sujet prise sa décision en la jugeant (selon un critérium déterminé) comme probablement correcte ou probablement incorrecte.

Par conséquent le modèle sur lequel cette méthode est basée peut être donc synthétisé dans le système d'équations sur lequel on base la théorie de la „signal detectability“, c'est-à-dire:

$$\begin{array}{lclcl} 1^{\circ} & \textit{Probabilité} & + & \textit{Probabilité} & = 1 \\ & (\text{identification correcte et valu-} & & (\text{identification correcte et valu-} & \\ & \text{tation correcte})^* (\text{appelée hit rate}) & & \text{tation incorrecte}) & \\ 2^{\circ} & \textit{Probabilité} & + & \textit{Probabilité} & = 1 \\ & (\text{identification incorrecte et} & & (\text{identification incorrecte et} & \\ & \text{valuation correcte}) (\text{appelée} & & \text{valuation incorrecte}) & \\ & \text{false alarm rate}) & & & \end{array}$$

Puisque la somme de chaque équation est 1, la prestation du sujet peut être définie seulement par deux paramètres, la „hit rate“ et la „false alarm rate“, et l'on peut tracer des courbes R O C (c'est-à-dire les courbes qui définissent la Receiver Operating Characteristic) telles que l'on a déterminées par les signaux sinusoïdaux.

Dans cette recherche nous avons étudié une modification de cette méthode à fin de faciliter son application.

En effet le procès peut être pratiquement réduit à un seul stade; la décision devient pour ainsi dire- automatique parce que le message est compris dans une liste étroite connue par le sujet. Par conséquent, le sujet ne doit pas évaluer sa décision, mais répéter le message tout simplement. Si les signaux qu'il a reçus ne sont pas suffisants

*Clinique Universitaire d'Otorhinolaryngologie de Ferrara (Italie), (Directeur: Prof. E. Bocca).

pour lui permettre de les identifier avec un des mots de la liste, on obtiendra une réponse négative.

Le modèle relatif est donc différent du précédent et peut être réduit à une seule équation:

$$\begin{array}{l} \text{Probabilité} \\ \text{(réponse correspon-} \\ \text{dant à l'original)} \end{array} + \begin{array}{l} \text{Probabilité} \\ \text{(réponse qui ne} \\ \text{correspond pas} \\ \text{à l'original)} \end{array} + \begin{array}{l} \text{Probabilité} \\ \text{(pas de réponse)} \end{array} = 1$$

Même dans ce cas il est suffisant d'utiliser deux membres de l'équation pour définir la réponse du sujet exactement.

Trois Collègues de notre Clinique ont été soumis à cette expérience.

Les messages qui avaient été enregistrés sur un ruban magnétique, étaient envoyés à un mixer auquel un bruit blanc arrivait aussi avec un niveau d'intensité global de 40 dB S. P. L.. Deux niveaux du rapport Signal/Bruit ont été employés pour l'épreuve, et précisément — 9 dB et — 6 dB.

Pour la préparation des messages nous avons premièrement rédigé une liste de 50 mots spondés que nous avons choisis parmi ceux que l'on emploie le plus communément.

Nous avons distribué ces 50 mots-clefs, trois fois, en 50 phrases; chaque phrase comprenait donc 6 syllabes des mots-clefs et 2—3 syllabes représentées par les articles, prépositions, etc., pour un total qui n'a jamais dépassé les 9 syllabes.

Les phrases étaient constituées par de simples expressions énonciatives de type prédicatif, avec quelques cas de type transpositif ou de type négatif.

A côté de ce matériel — selon le même critérium — nous avons aussi rédigé 50 nouvelles phrases, dans lesquelles les mots-clefs étaient liés les uns aux autres avec des probabilités conditionnelles très basses.

Dans ces conditions les phrases rappelaient le langage de certains écrivains surréalistes (par exemple „la lune boit l'enfant“).

Pendant l'audition (par les écouteurs) le sujet avait devant lui la liste des 50 mots qu'il avait mémorisés; l'intervalle entre une phrase et l'autre était prolongé jusqu'à ce que le sujet donnait sa réponse.

Les réponses pouvaient être divisées en trois types:

- a) Réponses correctes. (Nous avons considéré comme correctes même les répétitions contenant de petites fautes de détail, fautes qui sont presque constamment représentées par la substitution de l'article déterminé avec l'indéterminé).
- b) Réponses dans lesquelles un mot-clef avait été remplacé par un autre.
- c) Pas de réponse.

Les résultats paraissent dans le schéma suivant (Fig. 1), qui reporte la probabilité de la réponse: sur l'ordonnée la réponse a) (c'est-à-dire, correcte), et sur l'abscisse la réponse c) (c'est-à-dire, pas de réponse).

Les résultats aux deux niveaux de rapport Signal/Bruit sont indiqués pour chaque sujet.

Les symboles pleins marquent les phrases qui ont un commun sens réel, les symboles vides marquent au contraire les phrases dans lesquelles les probabilités conditionnelles des liaisons entre les mots-clefs étaient très basses.

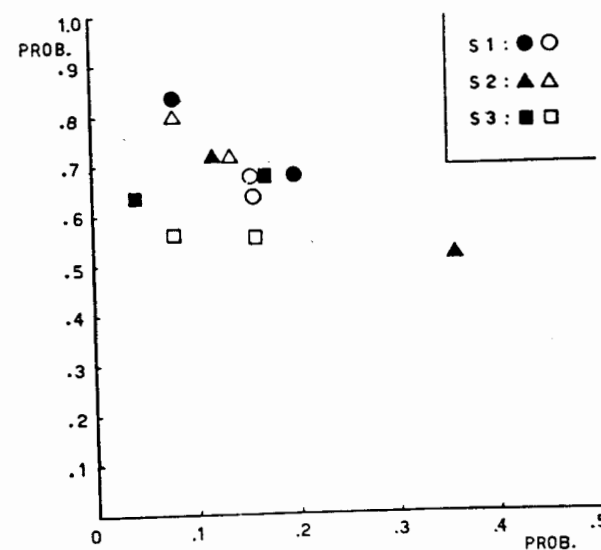


Fig. 1.

Ainsi que le schéma montre, dans leur ensemble les points paraissent très rapprochés et indiquent — par là — une bonne univocité du critérium de décision. Même l'analyse statistique, que nous avons conduite en comparant deux à deux les différentes réponses de chaque sujet par le test de l'exacte probabilité (c'est-à-dire le test de la table à quatre cases), n'a pas révélé de différences statistiquement significatives.

Une dernière remarque regarde précisément la différence entre les réponses obtenues par les phrases à sens réel et celles que nous pourrions définir des phrases „dépourvues de sens commun“. Cette donnée indique qu'une évaluation des réponses basée sur la méthodique que nous avons proposée, peut faire abstraction du type de message vocal employé.

En effet les résultats ne démontrent pas des différences substantielles entre eux, soit qu'il s'agisse de mots uniques, soit qu'il s'agisse des mêmes mots structurés en phrases et liés entre eux par des différentes probabilités conditionnelles.

ARTIFICIAL DISTORTION OF WORDS

MIHAJLO PRAŽIĆ

Hearing is a phenomenon which is composed of different elements, anatomical, biochemical, elektroneurophysiological, linguistic, phonetic and statistical. Therefore one can study hearing as a process either from a medico-audiological, psychological, or linguistic standpoint, which means, that the results from these rather different groups of sciences must be somewhat different, or, at least they cannot be identical. Medical audiology has developed tonal liminar audiometry, tonal supraliminar audiometry and with all these techniques and tests one can obtain precise results about the lesions of hearing from the outer part of the ear up to the cochlea with all its structures.

The attempts to draw out with supraliminar tonal audiometrical tests pure neural, or even central lesions of hearing, in spite of many elaborate tests could not find broader application in the audiological diagnosis.

In such a situation the audiologists turned to speech audiometry with the task of elaborating tests for the ascertainment of central disorders and lesions of hearing. In the last 10 years many distorted speech audiometrical tests have been elaborated. With the distortions of speech elements (the distortion was performed in different ways and principles) on the ground of full or lessened discrimination effect one should reach the diagnosis of the central lesions of hearing.

All these tests in spite of the fact, that they were theoretically ingeniously elaborated, have too many uncertain elements to be applicable in routine audiological examination. The main reason lies in the complex structure of the phenomenon of hearing. Human communication takes place with speech. But, what is speech?

Into the ear of the person who is listening do not enter either sentences and words or syllables, from which every single word is made up. Into the ear during communication enter only sequences of air vibrations or sequences of variations in the air-pressure. From the phonatory organ of the person come not sentences or words, but sequences of variations in the air-pressure. In the cochlea take place the transformation of the mechanical movements into electrical potentials, and a detailed and definitive frequency and amplitude analysis. The sequences of the impulses of the equivalent pass through the filaments of the acoustic nerve to the brain, where the definitive stage of the hearing phenomenon, the decoding will take place. The decoding is the linguistic and statistic process for the ascertainment and identification of the

corresponding acoustic picture which as an engram lies as a deposit in the "memory". In the same time interval take place the confrontation and identification of the sequences which arrive, with the adequate and corresponding acoustic engrams. But these sequences consist only of the series of impulses which flow through the acoustic nerve. As yet we know neither the total number of informations for a certain acoustic engram, nor the minimal quantity of informations on the ground of which an adequate identification will take place in the process of decoding. This minimal number of informations depends on the presence of many factors which are superimposed in this process such as noise etc.

Every distorted word, from the theoretical standpoint, should have its frequency composition (and in some tests also the time factor) so much distorted that the sequences of informations which reach the decoding process will not correspond well to the adequate normal engram.

It is obvious that during the presentation of such distorted words the whole linguistic and statistic procedure must be activated under extremely unfavourable conditions in order to reach the adequate engram in the process of decoding. Therefore the whole diagnostic procedure in the audiological examination of the lesion of hearing of presumably central origin takes place in the area of psychological, linguistic, phonetic and statistic elaboration and in a very small part in the medico-audiological ascertaining of the nature and the degree of the lesion of hearing.

The application of speech audiometrical tests for the topical examination of the lesion of hearing especially in the area of the cochlea has so far been only partly elaborated and has not shown clear results. Lafon in his phonetical testing and measuring, of hearing besides tests intended only for the diagnostical procedure for the ascertainment of the central types of lesion of hearing, elaborated also a test for the ascertainment of lesions of hearing located in the cochlea. In this test he uses the same linguistic and mathematical procedure as in the test for central lesions of hearing, and that cannot be the right way for this task.

It is a fact that with tonal audiometrical tests and methods it is impossible to ascertain the finer differentiation of the perceptive lesions of hearing of the sensory type. It is evident that all perceptive lesions of hearing of the sensory type cannot have the same pathological origin and the same anatomical background. Owing to the complex mechanism of the frequency and amplitude analysis which takes place in all the structures of the cochlea and not only in the sensory cells of the Organ of Corti, one should try to divide the broad group of sensory lesions of hearing into some special types and audiological forms, but during this procedure one should take maximum care not to disturb the final stage of the phenomenon of hearing. Only under such circumstances could one be sure that the phenomenon of hearing as a whole will take place under normal and not artificially aggravated conditions. All tests with distorted speech elements are elaborated either with frequency distortions or with the superposition of a changed time factor. Therefore the process of the frequency and amplitude analysis of a quite new speech element takes place in

the cochlea, and sequences which very little resemble the original, undistorted word run to the brain. For this reason the decoding process will take place under ad maximum aggravated conditions of linguistic and statistical elaboration, and the final identification will take place with or without success, but we cannot state or measure

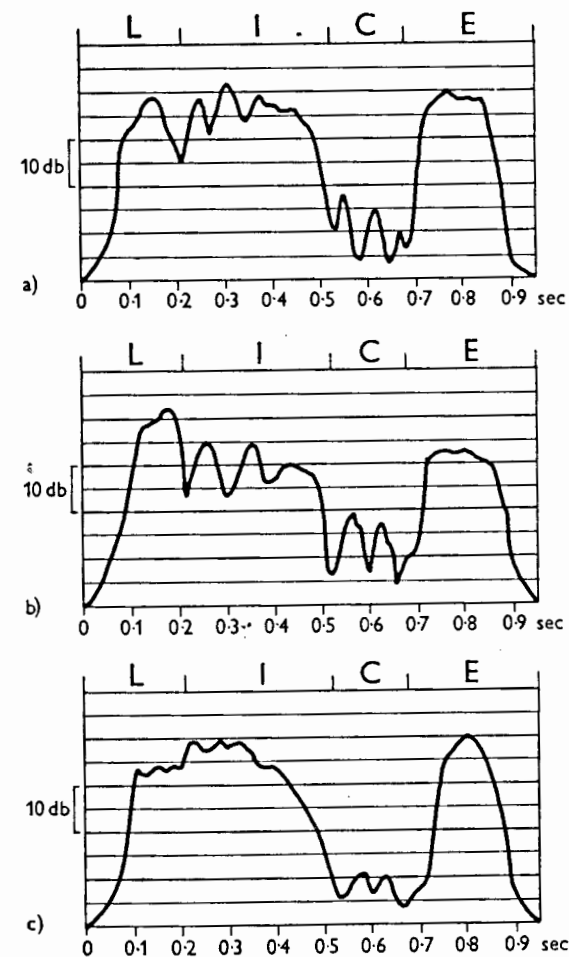


Fig. 1.

with non-objective methods all forces used in this final act of hearing. The examination of hearing with distorted speech elements with the aim of ascertaining finer forms in the group of sensory lesions of hearing can be successfully elaborated only if the distortion can be performed in a way which would attack only the analytical process in the cochlea but not the later stage of the decoding, which means that the distortion must be achieved in such a way that the sequences as a whole in passing from the cochlea to the brain must be ad maximum similar and adequate to the sequences of the normal presentation of the same speech elements. The frequency

distortion with filters changes too much the acoustic material which enters the cochlea. Therefore quite another process of analysis of this speech element must follow. It is obvious that every distortion of the frequency qualities of the speech elements at least for the purpose of studying the function of the cochlea must be without sufficient effect.

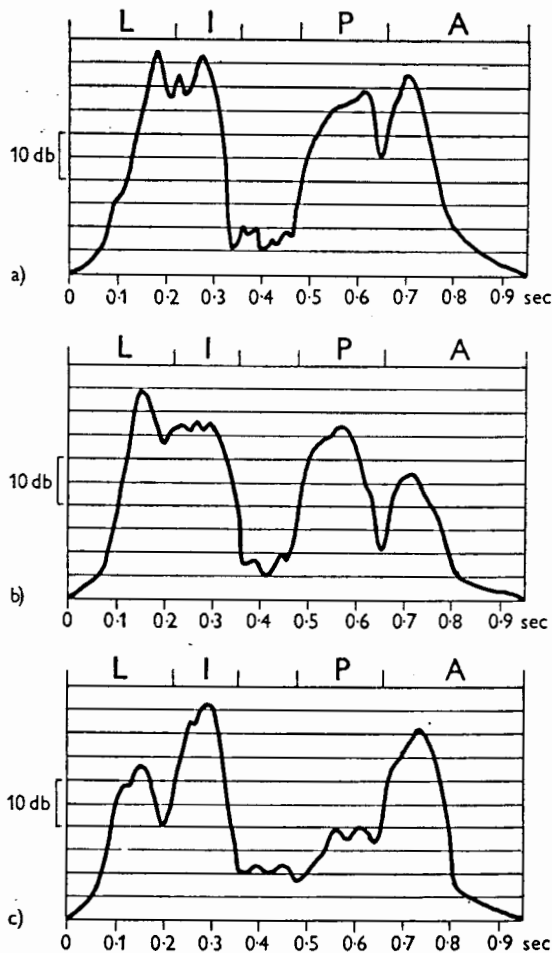


Fig. 2.

Another as yet not applied principle of distortion is, to perform the distortion only of amplitude elements and only of some of these which will be strictly selected in advance. If one can carry out this distortion precisely and measure it exactly, it should be possible to work with precise units of measurement such as the decibel.

I elaborated such a distortion test in which I lowered by 10 decibels the amplitude values of either vowels or consonants in the speech elements. With a special technique this procedure can be performed precisely, and the tape-recorder in the Brüel Kjaer logarithmic writer shows exactly if the procedure is performed successfully or not.

As one can see on the following figures 1, 2, 3, 4 and 5, every word has a certain envelope with frequency and amplitude elements for every phoneme. After a weakening either of vowels or of consonants the envelope shows that this weakening amounts to exactly 10 decibels either for consonants or for vowels. If we reproduce these

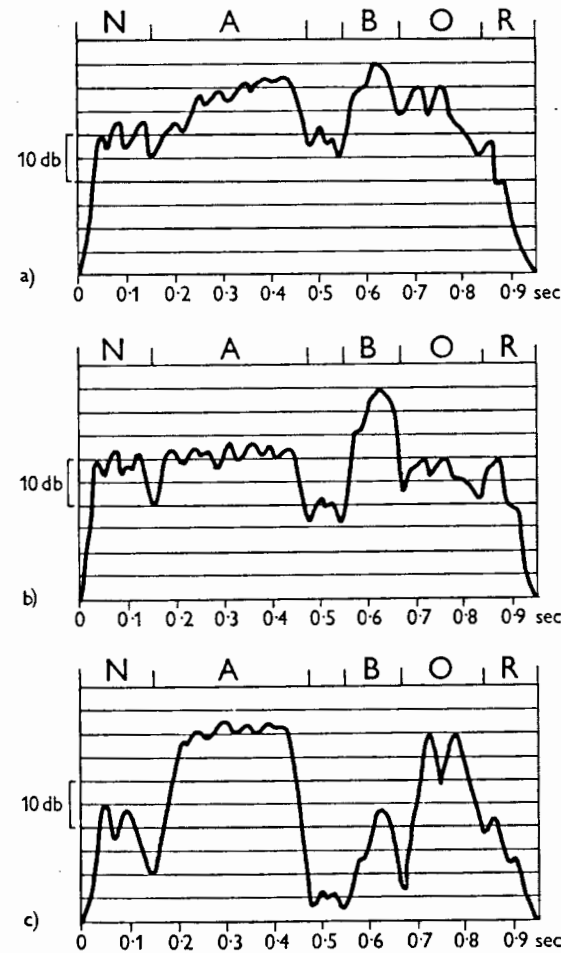


Fig. 3.

weakened words through a loud-speaker in the free field, the intelligibility of these artificially distorted words will be for the person with normal hearing only somewhat lessened. But, if we present these distorted words to patients with reduced hearing we shall see that there is an evident difference between conductive and perceptive losses of hearing. In all conductive cases the intelligibility of such distorted words with weakening either of vowels or of consonants presented at different supraliminal levels is the same as with persons of normal hearing. But with the patients with perceptive loss of hearing there are 2 different groups. In the one group the patients show mar-

edly lessened intelligibility of the distorted words at all higher supraliminal levels. In the other group the patients show the same intelligibility as the patients with

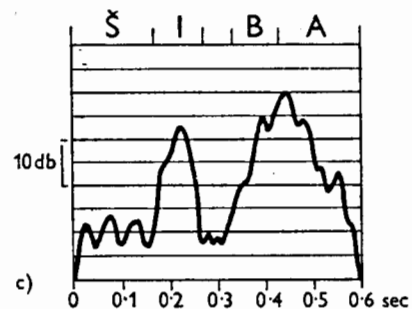
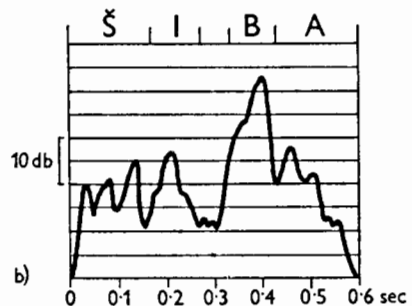
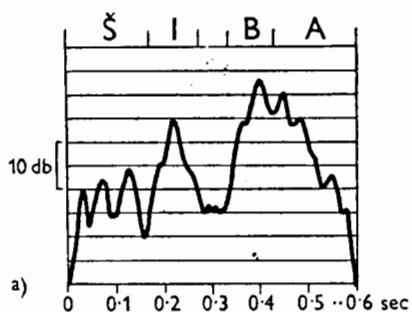


Fig. 4.

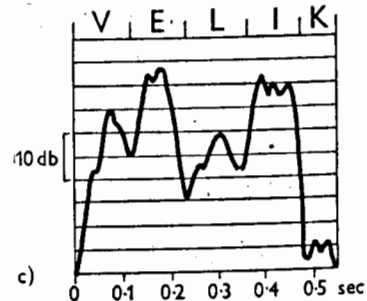
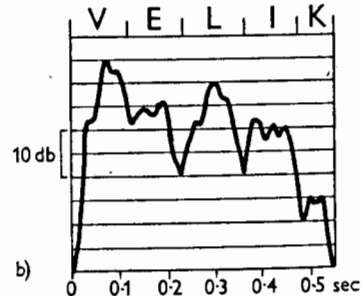
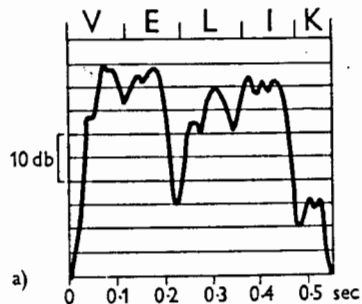


Fig. 5.

conductive loss of hearing. These results show that it is possible with the distortion test of weakening either of vowels or of consonants to obtain some new details on the audiological picture of perceptive losses of hearing.

PHONETIC SIGNALS OF SYLLABLE BOUNDARIES

ERNST PULGRAM*

At the last Phonetics Congress, in Münster, I suggested that the syllable should be regarded as a linguistic unit so constituted as to contain one and no more than one vowel, with the consonants preceding the vowel (onset) and following it (coda) to be determined by applying certain heuristic rules based upon the distributional properties and constraints prevailing in a given language.¹ (This presupposes a definition of vowel and consonant that is not circularly derived from their role in the syllable; such a definition on purely distributional grounds is available.)² I still hold this view in essence, although I have modified it in some details, thanks in part to the remarks made by colleagues in Münster, notably Messrs Fliflet, Trim, Hammarström, and Miss Sivertsen. A publication will soon be forthcoming, but I shall be grateful for further comments.

Rather than speaking of the details which I have altered somewhat since 1964, I shall today address myself to the problem of the phonetic signals that mark the syllable boundaries fixed by phonological-phonotactic means.

Most phoneticians would agree that the search for the universal phonetic traits signaling the syllable boundary has been less than successful. To be sure, it was found that various articulatory and acoustic events accompanied what both linguists and laymen would call a syllabic break in a given utterance; but none of these events, as far as I know, has turned out to be the indispensable signal that invariably marks a syllabic break and that never marks anything but a syllabic break. In other words, such signals as have been observed are not unambiguous and specific markers of syllable boundaries.

In part, I believe, this is so because, even though the syllable itself is a linguistic universal (every layman knows of it, and every linguist operates with it), there is no universal syllabic shape, and there is no universal event in the stream of speech whose

* University of Michigan, U.S.A.

¹ Ernst Pulgram, Consonant cluster, consonant sequence, and the syllable, *Phonetica* 13 (1965) 76—80, (discussion 80—81).

² See J. D. O'Connor—J. L. M. Trim, Vowel, consonant, and syllable — a phonological definition, *Word* 9 (1953) 103—122.

occurrence says 'Syllable boundary'. Every language has its particular distributional phonological constraints and phonetic signals at syllable boundaries.

Some years ago, that excellent phonetician, who unfortunately died in the midst of her career, Marguerite Durand, wondered whether the nonsense string *skuks* /skuks/ was one or more than one syllable, and she attempted to find an answer by phonetic means.³ But she was posing, I fear, a pseudo-problem to which there is no phonetic answer. To the speaker of English or German, /skuks/ is monosyllabic: it has one vowel, an onset /sk/, and a coda /ks/. But to a speaker of Spanish, who can have neither an onset /sk/ nor a coda /ks/ since the first does not appear postpausally and the second not prepausally, the word may well seem trisyllabic, and since he so hears it he also provides in speaking the requisite vowels and pronounces it /es-kuk-se/; in fact, if one took the position that even word-medial /k-s/ is un-Spanish because it places /k/ in coda position while Spanish does not have a prepausal /k/, one would have to posit four syllables, namely, /es-ku-ke-se/ — and tetrasyllabic it may sound to, and be pronounced by, speakers of Spanish.⁴ Consequently, the search for phonetic clues betraying the number and boundaries of syllables in *skuks* cannot but be, and has been in many strings of allophones thus examined, fruitless, or at least distressingly inconclusive.

There is yet another, and more impeditive, obstacle to the discovery of phonetic syllable boundary markers, namely, their optionality. I should affirm that the establishment of syllable boundaries on the phonological level by phonotactic means delivers the points where phonetic syllable boundary signals *MAY* occur; but they *NEED NOT* occur. That is to say, there need occur none of the typically unit-final or unit-initial allophones, no pause, no silence, no glottal stop, no clue whatever whereby a given language customarily signals segmental boundaries within the stream of speech. Clearly, therefore, this is one more reason why informants and linguists may disagree not only on the place of the syllabic boundary, but even on its phonetic reality: it simply is not always realized. One knows how many syllables there must be in a piece of discourse: as many as there are vowels; so one naturally wishes to find out how they are bounded against one another. Stating this boundary on the phonological-phonotactic level means no more than saying: This is the place where a syl-

lable boundary is going to be signaled — *IF* it is going to be signaled. But since it not always is, the expectation to discover a specific phonetic event that invariably takes place at every syllable boundary in all languages cannot but be disappointed.

It is this absence of *consistent* phonetic signaling which has prompted some linguists to deny linguistic status and reality to the syllable altogether. Considering, however, the universal usefulness and employment of the syllable in synchronic and diachronic linguistics, one is tempted to paraphrase Voltaire's dictum about God and avow, "Si la syllabe n'existait pas, il faudrait l'inventer."

But the notion of the optionality of something that is claimed to be a universal linguistic unit, demands a word of explanation.

According to what I have said, one might be inclined to conclude that the syllable boundary has two phonetic realizations: an appropriate segmenting signal such as the language under scrutiny provides, and zero. I am, however, reluctant to put it this way because it evokes the relation of phoneme and allophone, or morpheme and allomorph, and would undoubtedly lead to the positing of syllabeme and allosyllable. But that pair of terms is superfluous and misleading because the syllable is not an emic unit in the sense that it serves a distinctive function. (Of course, a syllable may coincide with a distinctive unit if it happens to be coextensive with a phoneme or a morpheme.) Syllabizing an utterance on the phonological-phonotactic level delivers therefore a structural statement on syllables and their shapes in a given language; syllabizing on the phonetic level is something that the speaker does if he so chooses, and that the listener may possibly 'hear' even if the speaker has not provided a signal. But the only reason why the speaker so chooses, or the listener so 'hears', is his desire to divide the utterance into syllables: nothing on any communicative or semantic level of linguistic analysis is thereby accomplished, no functional aim is pursued or attained (except in cases where insistence on syllabation disambiguates a message, e.g., "I said *pee-ring* not *pee-ling*"). Unlike other linguistic units, the syllable is its own and only purpose. And I venture to suggest that precisely this non-functional character of the syllable, this basic meaninglessness of it in an utterance, is what makes its phonetic realization optional, causes syllable boundaries not always to be marked phonetically where phonological-phonotactic analysis places them, and has rendered the signals of syllable boundaries so elusive in theoretical and experimental phonetics.

³ Marguerite Durand, La syllabe: ses définitions, sa nature, *Orbis* 3 (1954), 527—533.

⁴ As a matter of fact, words with medial /ks/ are rare in Spanish, and usually borrowed, like *axioma* /ak-sjo-ma/—which in popular speech would in any event be changed to /a-ke-sjo-ma/. Setting up a syllabation /a-ksjo-ma/, with syllable-initial /ks/, on the evidence of words composed of *xilo-* /ksilo/ and another morpheme, is not helpful because those words themselves are equally subject to being classified as not completely naturalized loans, and indeed to being pronounced, after naturalization, as /silo/. I am certain that whenever one is forced into a syllabation that runs counter to the distributional requirements of the language, one is faced with a non-assimilated loan (which includes learned words), or an obsolescent or ascendent form: cf. Ernst Pulgram, French /ə/: statics and dynamics of linguistic subcodes, *Lingua* 10 (1961) 302—325.

THE AUDITORY PERCEPTION OF DISTINCTIVE FEATURES IN A SIX YEAR OLD CHILD

WALBURGA VON RAFFLER ENGEL*

While I was pulling together the data derived from the bit of research I am going to present here, I came across an article¹ that strikingly supported my findings. The virtual identity of the author's conclusions and mine may be considered significant in that we started from different premises. Prof. Wickelgren is a psychologist and I am a linguist. Moreover, his research and mine had different purposes. Prof. Wickelgren tested errors arising in short-term memory and I tested errors stemming from spontaneous confusion. The subjects too were different. Prof. Wickelgren was dealing with adults while I was concerned with child language.

What had prompted my research was the historical concern of why phonological change—both assimilative and dissimilative, contiguous as well as discontinuous—operates prevalently on distinctive features.² In diachronic linguistics the phoneme may conveniently be considered as a unit, in accordance with the reconstructions of Prof. Hoenigswald.³ But too little is understood of the elements that make up the whole,⁴ even though, to my mind, the simultaneous presence of discrete features does not in itself detract from the unitary character of the whole.⁵ In all probability Prof. Wickelgren was approaching the problem from a synchronic point of view.

His "results suggest that a consonant is coded in short-term memory, not as a unit but as a set of distinctive features each of which may be forgotten at least-semi-independently."⁶ I, too, concluded that a consonant is coded as a set of distinctive

* Vanderbilt University, U.S.A.

¹ Wayne A. Wickelgren, Distinctive features and errors in short-time memory for English consonants. „*Journal of the Acoustical Society of America*“ 39, 2 (February 1966), 388—98.

² I am using this term in its broader meaning of phonological component rather than in the specific Jakobson-Halle terminology.

³ Cf. Henry M. Hoenigswald, *Language change and linguistic reconstruction*, U. of Chicago Press, 1960.

⁴ To give only one example that a componential analysis is philologically and structurally sound: during the third century B.C., when Greek loses prevocalic aspiration, it also changes its aspirates *ph* and *th* to the fricatives *f* and *θ*.

⁵ From another standpoint see the comment on the Householder-Chomsky controversy by Marvin D. Carmony, *Surface and deep phonology*, „*Journal of Linguistics*“, 2,2 (October 1966), 208.

⁶ *Op. cit.* Note 1, p. 388.

features. Each of these may be emphasized individually or connectedly, and induce the total or partial oblivion of the others.

At the present state of research we cannot posit the exact hierarchies of distinctive features—within each phoneme, paradigmatically, as opposed to all other phonemes, and, syntagmatically, as determined by the influence of preceding and following phonemes and in every prosodic situation.

Unfortunately I did not have time to consider the important related issue of consonant perception *per se*, whether the perception of a consonant is fully or partially conditioned by onset and offset of the surrounding vowels.⁷ Less vital to my hypothesis or Prof. Wickelgren's is the running controversy over the motor theory of speech perception.⁸

The subject whom I have been observing is my son. He is now six years and ten months old, has completed the first grade of an American school and can read with ease any book designed for children of his age or slightly older, but he does not seem to enjoy reading. He can write English fairly well and occasionally, of his own will, composes simple letters to friends and relatives.

At school, after an initial period of apathy towards the study of his language, he became an excellent speller. One of his favorite pastimes is asking his parents to give him words to spell, and of this he never tires. He also enjoys watching spelling bee programs on television. From infancy on, my son seemed to be more ear than eye minded.⁹

He started the first grade at six years and two weeks of age and was taught "phonics" using the regular English orthography, learning letter after letter. The teacher explained how each graphemic unit corresponded to a phonetic expression. My son considered homonyms like *meat* and *meet* as much fun as playing with a puzzle. Whatever misspelling has so far occurred to the child, a vowel was always confused with some other vowel and never with a consonant. The perception of vowels is undoubtedly different from that of consonants, a fact which has been experimentally demonstrated by the Haskins group.¹⁰

Among the non-vocalic confusions of my son, a special status must be assigned to the *h* phoneme, which was the only sound which could be replaced by zero. The

⁷ Cf. A. A. Hill, *Suprasegmentals, prosodies, prosodemes*, "Language," 37, 4 (Oct.-Dec. 1961), 460: "As machines tell us, it is necessary to assume that only the onset and offset of consonants like (*θ*) and (*f*) in *three* and *free* or *myth* and *miff* are audible. Thus the real cue as to whether a word contains (*θ*) or (*f*) is the audible difference in the surrounding sounds."

⁸ Cf. Harlan Lane, The motor theory of speech perception: a critical review, „*Psychological Review*”, 72, 4 (1965), 275—309.

⁹ Cf. Walburga von Raffler Engel, *Il prelinguaggio infantile*. Paideia 1964, pp. 44, par. 33, 70—71, par. 63.

¹⁰ Cf. D. B. Fry, Arthur S. Abramson, Peter D. Eimas, and Alvin M. Liberman, The identification and discrimination of synthesis vowels, "Language and Speech," 5, 4, (Oct.—Dec. 1962), 171—189.

same phenomenon, as a matter of fact, happens all the time when native speakers of a Romance language are learning a Germanic language. I played at spelling new words with the child between the third and sixth month of his first school year. At that time, he was speaking English like any other American youngster, but it must be remembered that my son is an Italian English bilingual and the English *h/θ* correlation might have been an interference from Italian. During the period of testing via word spelling games, I have, however, observed no other bilingual phonic interference. As a matter of fact during the subject's earlier speaking period, the difficult phoneme *r* was replaced by *w* in English and by *l* in Italian. It thus remains unclear how to interpret utterances like: "Mother, I know two more words that sound alike and are spelled differently: *his* and *is*" (February 18, 1967).

Despite his normal command of spoken English, a little note my son jotted down once about his parents revealed orthographic errors quite similar to the mistakes vocalized by children with speech disturbances: "I weley (really) like them wele [very] much" (April 23, 1967). Voicing is shared by *r*, *l*, *w*, *v*. As a matter of fact no confusion ever arose in his speech between voiced and unvoiced sounds. As has been observed by several scholars, the vibration of the vocal cords seems to be a prime element in the hierarchy of phonological components. Acoustically, *r*, *l*, *w* are all of low frequency. My son confused also *y* and *l*. Obviously, the semi-vowels formed a distinct category. The intrusion of *v* needs further investigation and may be due to nothing more than graphemic similarity.

Among the consonants proper no breakdown by category could be established. Within the stops, *key* and *tea* were considered homonyms and *b* and *d* were confused; but at the same time, the final consonant in *handkerchief* became a stop in *handkerchip*. Here the labiality outweighed the fricativeness. In other words of his, where he substituted *f* for *θ*, the acoustical high frequency component of fricativeness seemed to override both place of articulation and tongue position. From a statement the youngster made during the second half of the first grade (March 20, 1967), it would appear that acoustic perception plays a major role. He complained about the complexities of spelling: "They all sound alike, and you never know if the word is spelled with an *s* or a *c*, like *see* and *cents*. And *c* and *k*: *cat* and *kitten*. Or sometimes they spell with a *k*, and sometimes with a *t*. How do you spell *cake*, with a *k* or with a *t*?"

The collection of my son's spelling errors leads not only to the conclusion that the phoneme is perceived as a set of discrete components, but also that the word or, for that matter, the morpheme need not be conceived of as the unit of perception. Throughout the entire first grade, I was on the lookout for interferences from the plerematic level, and so far, have been able to detect only one instance: "I know a word that ends with *i*, *cab*." When I asked the child to write it down, he wrote *taxi*. The latter confusion points to a "global conceptualisation" on the semantic level.¹¹

¹¹ G. Walburga von Raffler Engel, *The concept of sets in a bilingual child, paper to be presented at the Xth International Congress of Linguists, Bucharest 1967.*

Coming as an afterthought, the phoneme, thus, still seems to perform an essential purpose, but we just do not know enough about the phoneme to determine its function in the missing link between physical perception and psychic interpretation.

DISCUSSION

Becker:

Wir konnten ebenfalls ähnliche, jedoch keine völlig übereinstimmenden Fehlleistungen bei Dyslalikern und Dyslektikern beobachten. Dyslektikern unterlaufen beim Schreiben z. B. häufig Vokalauslassungen, Dyslalikern beim Sprechen nicht bzw. selten. Die Fehler bei Konsonanten unterscheiden sich auch.

Beispielwort Knöpfe = Dyslaliker spricht: Nöppe
 = Dyslektiker schreibt: Köfe
Stube = Dyslaliker spricht: Tube
 Dyslektiker schreibt: Schube oder Schub

Die Fehler scheinen bei Dyslalikern stärker durch Artikulationsschwierigkeiten, bei Dyslektikern mehr durch phonematische Differenzierungsschwäche bedingt. Nach unseren Beobachtungen werden Explosiv- u. Nasallaute von allen Kindern am schlechtesten phonematisch differenziert. (vgl. Beltjukow u. Theiner). Wir stimmen mit der Autorin überein, bei Dyslektikern der phonematischen Differenzierungsfähigkeit mehr Aufmerksamkeit zu schenken.

ИНТОНАЦИЯ ОБРАЩЕНИЯ

(на материале английского языка)

Л. Д. РЕВТОВА*

Обращение в английском языке характеризуется, как известно, большим разнообразием интонационных рисунков.

Основной целью работы и явилось установление собственно-коммуникативных видов обращения на основе восприятия их интонационных структур, а также раскрытие их семантики. Исследование позволило сделать и некоторые другие выводы об обращении, как коммуникативной единице английского языка.

ОГРАНИЧЕНИЯ И МЕТОДИКА ИССЛЕДОВАНИЯ

Аудиторскому анализу подвергалось свыше 300 фраз в контекстах диалогической речи, произнесенных дикторами-англичанами (художественное чтение, записи для учебных пособий и др.).

Установление семантики интонационной структуры обращения проводилось с учетом конкретной коммуникативной ситуации, так как она однозначно выявляет смысловое содержание интонационной структуры.

Вводятся понятия широкого и узкого контекста коммуникативной ситуации. Узкий контекст ситуации характеризуется тем, что лицо, к которому адресовано обращение, непосредственно участвует в разговоре. Широкий контекст ситуации предполагает лишь присутствие лица при разговоре, но не его непосредственное участие в нем.

Рассматривая интонационные структуры обращения, мы подвергли детальному анализу лишь два звена коммуникативной цепи речи — воспринимаемое и семантическое.

Аудиторский анализ прошел три этапа. В задачу первого этапа (5 аудиторов) входила фиксация воспринимаемых качеств интонационной структуры обращения в ситуации по схеме:

* ЛЭФ Минского государственного педагогического института иностранных языков, Минск.

И н д е к с	Слово-обращение в ситуации	Интонационные характеристики ¹⁾			
		Высотный уровень (В, Н, Ср)	Движение мелодики (\, /, \, /, —, —, \, /)	Степень ударения (у, б/у)	Темп произнесения (Б, С, М)
	Harry, something has gone wrong with my electric iron	? (Н)	? (V)	? (y)	? (С)

1) В — высокий, Н — низкий, Ср — средний, \ — простой нисходящий тон, / — простой восходящий тон, — — ровный тон, \, /, \, / — сложные тоны, у — ударный, б/у — безударный, В — быстрый, С — средний, М — медленный.

На втором этапе, пяти другим аудиторам предлагалось: установить правильность интонационного комплекса воспринимаемых качеств, указанных аудитором первого этапа анализа, и заполнить схему формальных признаков коммуникативной ситуации.

И н д.	Коммуникативная ситуация	Узкий контекст	Широкий контекст	Интонацион. комплекс. восприним. качеств	Пометка „да“, „нет“
15	Harry, something has gone wrong with my electric iron	? (+)	? (—)	Н, \, у, С	? (да)

Третий этап анализа ставил целью выявление семантики отдельных интонационных структур обращения вне ситуации. Аудиторы (5 лучших из 10 участвовавших в первых двух этапах анализа) должны были подчеркнуть воспринимаемые интонационные качества предложенного слова-обращения, установить правильность формальных признаков ситуации, описать ее и назвать основное семантическое значение обращения.

И н д.	Словообращение	Интонац. хар-ки				Признаки ситуации		Семантическое значен.	Дополнит. оттенки значения
		уровень	мелодика	ударен.	темп	широкий	узкий		
15	Harry	ВНСр	\, / ^, \	у бу	БОМ	+	— ? (да)	? (прямое обращение)	? (спокойное обращение)

Первый этап анализа установил следующие возможные комплексы интонационных характеристик обращения:

В \ у М; В \ / у М; В / \ у М; В / \ / у М; Н / у С;
Н \ / у С; Н — б/у Б; В / б/у Б.

Второй этап анализа определил следующие соответствия между интонационными структурами обращения и формальными признаками коммуникативной ситуации:

Участие в широком контексте	Участие в узком контексте	Соответствующие интонационные комплексы воспринимаемых качеств
—	—	В \ у М; В \ / у М; В / \ у М; В / \ / у М;
+	?	Н / у С; Н \ / у С;
+	+	Н-б/у Б; В-б/у Б;

В результате третьего этапа указаны следующие семантические характеристики обращения: обращение, зовущее издали; призывающее прийти; непосредственное обращение; прямое обращение; слово-обращение, упомянутое для вежливости, но не для зова; слово-обращение, употребленное для стиля; традиционное обращение, лишенное своего основного назначения и подобные.

К дополнительным оттенкам значения были отнесены: небольшое раздражение, нетерпение, спокойствие, игривость, безразличие.

Исследование интонации обращения позволило различать три вида обращения в английском языке: *вокативный*, *апеллятивный* и *неапеллятивный*.

DISCUSSION

Анна:

There seems to be a mixture of levels involved in the paper. The distinction between different styles of discourse and intonation patterns should be rigidly maintained. It seems strange that examples were not provided and that the contextualisation of those given was not clarified. In my opinion the division into three categories is not sufficient to account for different registers or styles within English.

Артемов:

Интонационные исследования привлекают все больше число фонетистов во всех странах, особенно в нашей стране. Назрела необходимость или включить этот вопрос в повестку следующего конгресса или созвать симпозиум между 6-ым и 7-ым конгрессами.

INTERSUBJEKTIVE ÜBEREINSTIMMUNG VON INTONATIONSBEURTEILUNGEN

HELMUT RICHTER

Für den Erfolg des Abhörens sprachlicher Äußerungen gibt es zwei wichtige Kriterien: a) die Übereinstimmung der Abhörurteile mit Merkmalen, die dem Gegenstand der Beurteilung immanent sind, b) die Übereinstimmung der Urteile verschiedener Abhörer hinsichtlich bestimmter Merkmale (intersubjektive Übereinstimmung).¹

Die Problematik beider Kriterien kann statistisch als Frage nach *Korrelationen* bestimmt werden. Da die Abhörurteile nur selten als Werte einer quantitativen Variable gegeben sind, können die herkömmlichen korrelationsstatistischen Maßzahlen vielfach nur approximativ berechnet werden. Hier soll eine informationstheoretische Technik exemplifiziert werden, deren Ergebnisse den gleichen Aussagewert haben wie Korrelationsmaße, deren Anwendbarkeit aber vom Charakter der Variablen unabhängig ist. Es handelt sich um die im wesentlichen auf McGill zurückgehende *multivariate Informationsanalyse*.²

Die Daten der Exemplifikation beruhen auf einem Versuchskorpus, das primär zu Bestimmung des kommunikativen Gehalts verschiedener Intonationen von *ja* und *nein* — als Antworten auf eine Alternativfrage — benutzt wurde.³ Verf. sprach dafür die Wörter *ja* und *nein* unter systematischer Variation der Melodie und der Dauer der Melodieabschnitte auf Tonband. Kombinatorisch wurden alle Variationen zur Klasse 1 und 2 ohne Wiederholung aus den drei Elementen „gleichbleibend“ (*g*), „steigend“ (*s*) und „fallend“ (*f*) gebildet (*g, s, f, sg, fg; gs, fs; gf, sf*), wobei jede eingliedrige Versuchseinheit nach „kurz“ und „lang“, jede zweigliedrige nach „kurz-kurz“, „lang-kurz“, „kurz-lang“, „lang-lang“ variiert wurde. Insgesamt ergab dies für *ja* und für *nein* je 30 verschiedene Intonationen. Die 60 Versuchseinheiten wurden in einer Zufallsfolge dargeboten.

Es wäre nicht korrekt, in der gegebenen Beschreibung des Versuchskorpus ohne weiteres eine Angabe signalakustischer Eigenschaften zu erblicken. Die Beschreibung bezieht sich vielmehr auf die Ziele der Aktivität eines menschlichen Generators. Deshalb werden wir nicht von einer Variable „stimuli“, sondern von der Variable *Produktionen* sprechen.

¹ Vom Problem einer intrasubjektiven Übereinstimmung sehen wir hier ab.

² Vgl. Attneave F.: *Information theory in psychology*, New York 1959.

³ Vgl. Richter H.: *Zur Intonation der Bejahung und Verneinung im Hochdeutschen*; in *Sprache der Gegenwart*, Düsseldorf 1967.

Im Rahmen einer Übung an der Universität Köln hatten 5 Studenten das Korpus abzuhören. Zur Beurteilung standen die gleichen Kategorien wie oben zu Verfügung. Das Experiment lieferte 5 zweidimensionale Häufigkeitsverteilungen über Wertkombinationen der Variable Produktionen (p) mit der Variable Reaktionen (r). Schichtet man nun diese fünf Verteilungen übereinander, so erhält man eine dreidimensionale Häufigkeitsverteilung über Wertkombinationen der Variablen, p, r und Versuchspersonen (s) mit den folgenden Summenverteilungen — zweidimensional: p/r , p/s , r/s ; eindimensional: p , r , s . Die verhältnismäßig geringe Zahl der Urteile je Vp erfordert bei unserem Material eine Zusammenfassung von Variablenwerten der Ausgangstabelle. Wenn wir Melodieverlauf und Dauer getrennt betrachten, können wir für den Melodieverlauf eine bivariate, für die Dauer eine trivariate Informationsanalyse durchführen.

Aus den Verteilungen werden folgende Entropiemaße direkt berechnet: die einfachen Entropien $H(p)$, $H(r)$ und $H(s)$, die zweidimensionalen Gesamtentropien $H(p, r)$, $H(p, s)$, $H(r, s)$ und die dreidimensionale Gesamtentropie $H(p, r, s)$.

Es gilt der Satz, daß bei Unabhängigkeit zweier Variablen p und r $H(p, r) = H(p) + H(r)$ bzw. $H(p, r, s) = H(p) + H(r) + H(s)$. Je stärker der Zusammenhang zwischen den Variablen ist, desto deutlicher wird die Gesamtentropie unter der Summe der einfachen Entropien bleiben. Die Differenz zwischen Summe der Einzelentropien und zugehöriger Gesamtentropie wird dadurch zu einem Maß des Zusammenhangs. Eine komplette trivariate Analyse liefert aus den dargestellten Grundmaßen durch Summen- und Differenzbildung über 20 weitere Maßzahlen, die weitgehend als Entsprechungen der einfachen, multiplen und partiellen Korrelationskoeffizienten zu deuten sind.

Wir beschränken uns hier auf eine Betrachtung von sog. *bedingten Entropien*. Während z. B. die einfache Entropie $H(r)$ angibt, welche Unsicherheit im Durchschnitt bei der Vorhersage einer bestimmten Reaktionskategorie besteht, gibt die bedingte Entropie $H_p(r) = H(p, r) - H(p)$ an, wie groß diese Unsicherheit noch ist, wenn bekannt ist, auf welche Produktion reagiert wurde; parallel dazu bezeichnet die bedingte Entropie $H_s(r)$ die durchschnittliche Unbestimmtheit einer Reaktionskategorie bei Bekanntheit der Vp. In einer trivariaten Analyse läßt sich schließlich berechnen, wie groß die Entropie von r noch ist, wenn Werte sowohl von p als auch von s gegeben sind: $H(p, r, s) - H(p, s) = H_{ps}(r)$.

In Analogie zur üblichen Definition der *relativen Entropie* als Quotient aus tatsächlicher Entropie und Maximalentropie wurden die bedingten Entropien aus dem Abhörexperiment auf die gleich 100 % gesetzte nicht-bedingte Entropie $H(r)$ bezogen; $\hat{H}_p(r) = 100 \cdot H_p(r)/H(r)$ usw.⁴ Wir erhielten das Bild der folgenden Tabelle bzw. graphischen Darstellung:

$\hat{H}(r)$	$\hat{H}(r)$	$\hat{H}_s(r)$	$\hat{H}_p(r)$	$\hat{H}_{ps}(r)$
103,5	100	97,2	38,2	
103,5	100	98,8	40,7	32,3

Bei den Melodieverläufen wie bei den Dauern wird die Unbestimmtheit des Abhörurteils durch Kenntnis der Vp nur wenig (um 2,8 % bzw. 1,2 %), durch Kenntnis der Produktion dagegen um etwa 60 % reduziert. Die Kenntnis der Vp zusätzlich zur

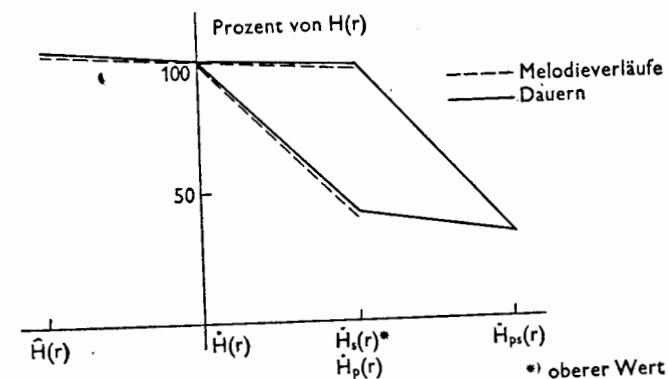


Abb. 1.

Kenntnis der Produktion (Dauern!) bedeutet keine wesentliche Erhöhung der Chancen, eine Reaktion vorherzusagen. Auf Grund dieser Ergebnisse unseres Erkundungsexperiments läßt sich die Hypothese aufstellen und in weiteren Versuchen zur Entscheidung bringen,⁵ daß individuelle Reaktionstendenzen bei der auditiven Beurteilung einfacher Intonationsmuster eine geringe Rolle spielen und die Nicht-Übereinstimmung zwischen Produktion und Reaktion in beträchtlichem Maße auf Verarbeitungstendenzen beruhen könnten, die den Mitgliedern einer Sprachgemeinschaft gemeinsam sind.

⁴ Um eine Verwechslung mit der üblichen relativen Entropie zu vermeiden, wurde auch die maximale Entropie $\hat{H}[r]$ auf $H[r]$ bezogen.

⁵ Die einschlägigen Informationsmaße können auch auf Signifikanz geprüft werden.

L'ASSIMILATION CONSONANTIQUE DE SONORITÉ EN FRANÇAIS: ÉTUDE ACOUSTIQUE ET PERCEPTUELLE

ANDRÉ RIGAULT*

C'est un fait bien connu que lorsque deux consonnes de sonorité différente sont en contact dans la chaîne parlée, il se produit une assimilation, l'une des consonnes imposant sa sonorité ou sa sourdité à l'autre. Les phonéticiens français sont toutefois d'accord que cette assimilation est complète au sein d'un mot lorsque les deux consonnes sont en contact permanent: *adscrite* [atskrit], mais partielle lorsque le contact résulte de la chute d'un *e* muet à l'intérieur d'un mot: *médecin* [medsɛ̃] ou de la rencontre fortuite de deux mots différents: *guide savant* [gid̥savã]. Ebauchées par P. Passy (*Les sons du français*, p. 113), développées par M. Grammont (*Traité pratique de prononciation française*, pp. 96—97), des règles subtiles de fonctionnement de l'assimilation consonantique de sonorité en français ont été énoncées en détail par P. Fouché (*Traité de prononciation française*, pp. XLVI—XLIX, 322, 324, 339, 349, 355, 366, etc.). Grammont et Fouché indiquent que, dans le cas d'assimilation partielle, la sourde sonorisée reste forte et que la sonore assourdie reste douce. Nous avons examiné de près de nombreux cas d'assimilation ayant lieu dans les différents contextes mentionnés ci-dessus afin de vérifier l'exactitude de ces règles. Cette communication devant être brève, nous nous limiterons ici à quelques cas.

Par exemple, l'assimilation de sonorité d'un [d] sous l'influence d'un [s] suivant a été étudiée dans les mots ou groupes suivants: *adscrite*, *médecin*, *tout de suite*, *guide savant*. La nature phonétique du [d] assimilé étant en cause, d'autres mots ou groupes contenant les séquences [d + z]: *razzia*, *Ladezan*, *tant de zèle*, *guide zoulou*, et [t + s]: *Batsère*, *Hauteserve*, *nous te servirons*, *fuite secrète*, ont également soumis à l'examen afin de servir de base de comparaison.

Ces mots et groupes de mots ont été insérés dans des phrases (ex: „c'est un médecin de renommée mondiale“) qui ont été lues, aussi naturellement que possible, par deux sujets français, nés et élevés respectivement dans la région parisienne et à Lille, et ayant une prononciation standard. Le corpus a été enregistré sur un magnétophone de haute qualité (Crown).

L'étude instrumentale a été faite, au point de vue articulatoire, à l'aide de l'aéromètre de S. Smith permettant de mesurer l'expiration buccale et l'expiration nasale,

* Université McGill, Montréal.

et d'un oscillographe, relié à un microphone laryngal, permettant d'obtenir un tracé de l'activité des cordes vocales. Au point de vue acoustique, l'étude a été menée à l'aide d'un sona-graph.

L'observation des tracés obtenus à l'aide de la combinaison aéromètre-oscillographe (figures 1, 2) montre que dans tous les cas, le /d/ assimilé présente les mêmes caractéristiques que le [t] de la séquence /t + s/: absence de vibrations laryngales, même tracé de l'activité buccale, mais par contre diffère radicalement du [d] de la séquence /d + z/.

L'examen des sonagrammes — auxquels sont joints des oscillogrammes de l'activité laryngale (figures 3, 4) fournit des observations analogues: absence de fondamental, même tracé d'amplitude pour le /d/ assimilé que pour le [t] de la séquence /t + s/.

Nous avons ensuite découpé la bande magnétique contenant les mots soumis à l'examen en isolant le groupe consonantique, précédé et suivi d'une voyelle, de façon à rendre le mot ou le groupe non identifiable: *médecin* /edsē/ *guide savant* /idsa/ etc. Les syllabes provenant de ces mots tronqués ont été jouées au magnétophone à cinq sujets français non-phonéticiens ayant reçu la consigne d'écrire aussi exactement que possible ce qu'ils allaient entendre.

Il est intéressant de noter que les sujets ont eu parfois de la difficulté à percevoir le lieu d'articulation de la consonne assimilée, /d/, par exemple, étant interprété dans quelques cas comme *p* ou *k*. Toutefois, puisqu'il s'agissait d'étudier la sonorité, nous n'avons tenu compte que de la sonorité indiquée dans la transcription, par les sujets, de la consonne assimilée.

Il ressort de cet examen que le /d/ assimilé est perçu dans presque tous les cas comme une sourde (figure 5). D'autres cas d'assimilation fournissent des résultats aussi nets, qu'il s'agisse d'assourdissement ou de sonorisation: dans /obsō/: *robe sombre*, le /b/ est perçu et interprété comme [p]; le /k/ de la séquence /akbo/: *paquebot*, comme un [g]; le /s/ de /asga/: *chasse gardée* comme un [z], et ceci par tous les sujets.

Dans quelques cas cependant, les sujets ne réagissent pas de façon aussi unanime. Ainsi, le /k/ de /ekdo/: *anecdote*, et de /akzo/: *sac jaune* est perçu comme [g] par trois sujets mais comme [k] par deux sujets. Chose plus curieuse, le /t/ de /ētvua/: *je viens te voir*, et le /d/ de /ēdvua/: *je viens de voir*, sont tous deux interprétés comme [d] par trois sujets et comme [t] par deux sujets.

Les résultats de ces tests, joints à l'observation des différents tracés (aérogrammes, oscillogrammes, sonagrammes) nous amènent à formuler les conclusions suivantes quant à l'assimilation consonantique de sonorité pratiquée par deux Français cultivés utilisant la prononciation standard:

1. La distinction de quatre catégories de consonnes, (a) sourdes-fortes [p t k f s ʃ], (b) sonores-douces [b d g v z ʒ], (c) sourdes-douces [b̥ d̥ g̥ v̥ z̥ ʒ̥], (d) sonores-fortes [p̣ ṭ ḳ f̣ ṣ ʃ̣] nous semble sujette à caution en français, au strict point de vue phonétique. Chez nos sujets, les catégories (c) et (d) n'existent pas.

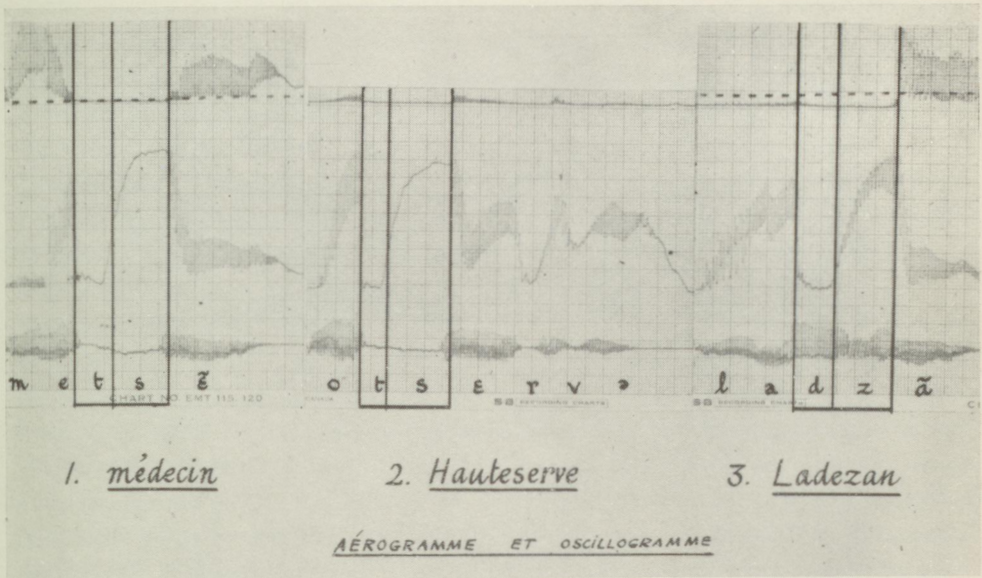


Fig. 1.

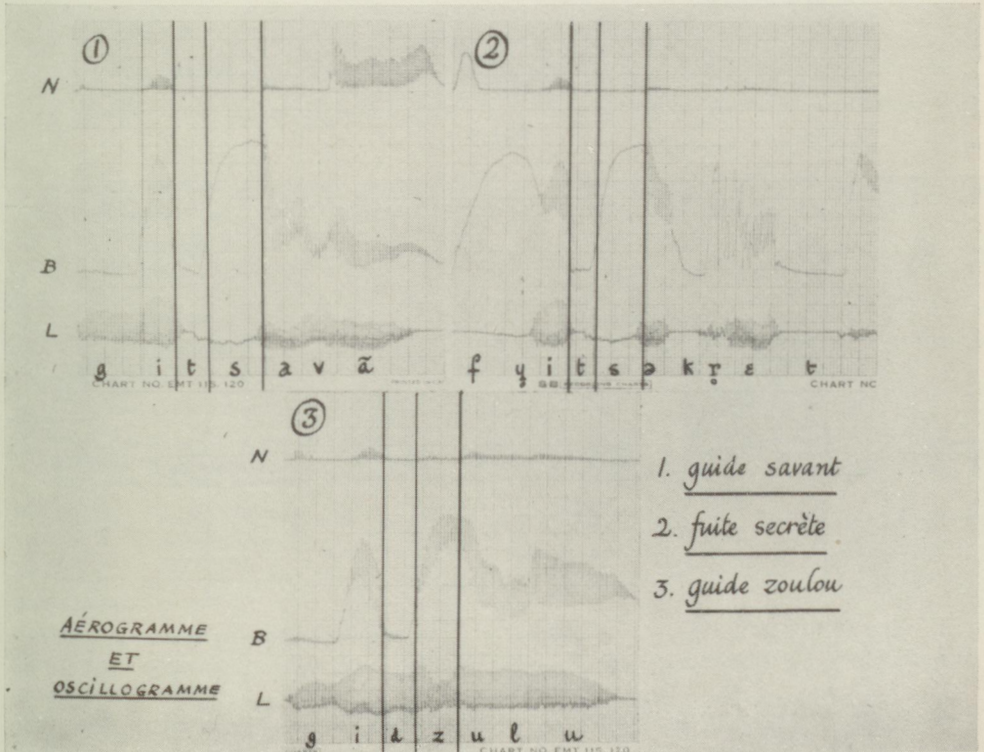


Fig. 2.

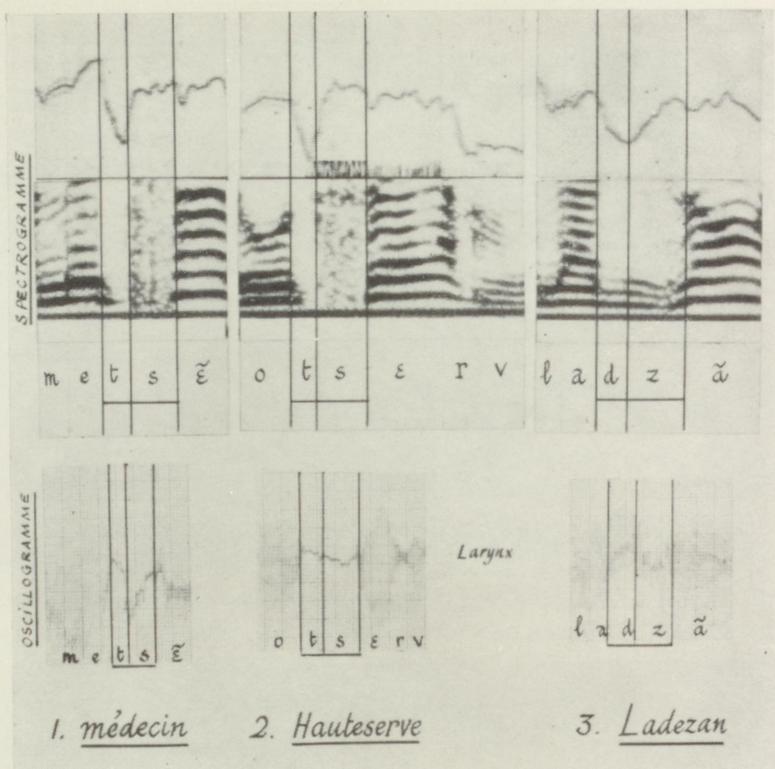


Fig. 3.

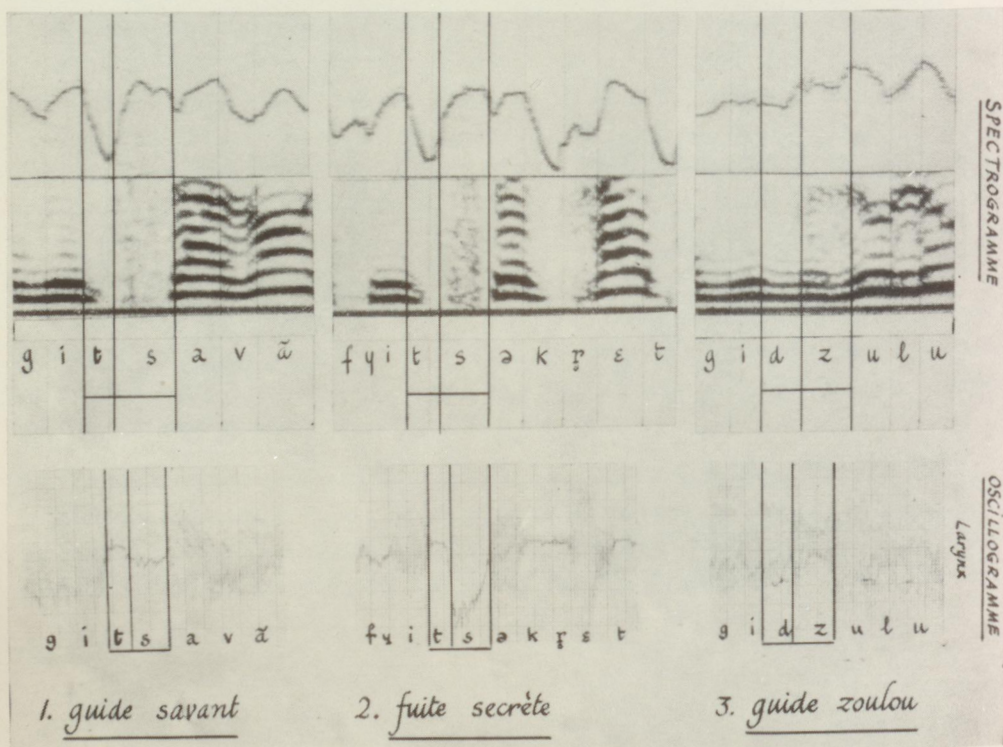


Fig. 4.

2. Par conséquent, la distinction établie entre assimilation complète et assimilation partielle nous paraît inutile puisque ne correspondant pas à la réalité des faits.
3. L'hypothèse émise par A. Martinet est confirmée: „En fait, il semblerait que dans des cas d'assimilation en contact, les traits distinctifs soient ceux qui s'imposent d'un phonème à un autre“ (*Economie des changements phonétiques*, p. 233). La sonorité — et non la force — est un trait distinctif en français.
4. Dans ces cas d'assimilation présentés sous forme de mots tronqués non-identifiables celui de /d + s/ exemple, il faut supposer ou bien que [t] est un allophone de /d/, ce qui est pour le moins difficile à admettre, ou plutôt qu'il y a neutralisation de la distinction sourde-sonore.
5. En prenant toujours comme exemple le cas de l'assourdissement de /d/ dans la séquence /d + s/, il faut admettre que les sujets interprètent différemment la consonne assimilée selon qu'elle se trouve dans une séquence phonique sans signification ou dans un monème identifiable. La même substance phonique [t] est perçue et interprétée comme /t/ dans les mots tronqués /edsē/, /idsa/, /otse/, /itsə/, alors qu'elle est interprétée comme /d/ dans les mots complets /medsē/: *médecin*; /gidsavā/: *guide savant*, mais comme /t/ dans /otserv/: *Hauteserve* et /fyitsəkret/: *fuite secrète*. Autrement dit, /d/ assourdi est perçu et interprété selon sa substance phonique dans un cas et selon sa forme linguistique dans l'autre.

	A	B	C	D	
				sourde	sonore
/d + s/	adscrite	/adskri/	[atskri]	4	1
	médecin	/edsē/	[etsē]	5	0
	tout de suite	/udsyi/	[utsyi]	4	1
	guide savant	/idsa/	[itsa]	5	0
/t + s/	Batsère	/atse/	[atsɛ]	5	0
	Hauteserve	/otse/	[otsɛ]	5	0
	nous te servirons	/utse/	[utsɛ]	3	2
	fuite secrète	/itsə/	[itsə]	5	0
/d + z/	razzia	/adzja/	[adzja]	0	5
	Ladezan	/adzā/	[adzā]	0	5
	tant de zèle	/ādze/	[ānzɛ]	0	5
	guide zoulou	/idzu/	[idzu]	0	5

A. Mots complets.

B. Syllabes provenant des mots tronqués.

C. Transcription phonétique de (B) d'après l'observation des spectrogrammes, aérogrammes et oscillogrammes.

D. Interprétation de la sonorité ou de la sourdité de /d/ assimilé, /t/ et /d/ par les cinq sujets.

Figure 5. Perception et interprétation de /d/ assimilé, /t/ et /d/ dans les mots tronqués.

Il n'y a donc pas correspondance entre la perception d'un signal acoustique à l'état brut et son interprétation dans la langue. En fait, l'identification des unités phonologiques ne peut s'opérer correctement que dans le cadre d'un syntagme signifiant.

DISCUSSION

Fischer—Jørgensen:

M. Thorsen (Copenhague) a fait une investigation instrumentale de l'assimilation de sonorité en français basée sur la prononciation de 5 sujets mâles (assez jeunes). Il a trouvé que l'assimilation de voix est normalement accompagnée d'une assimilation complète de la pression de l'air dans la cavité buccale, d'une assimilation complète de la pression des lèvres (par les labiales) et d'une assimilation incomplète de la durée (de la consonne et de la voyelle précédente). Voir Annual Report of the Institute of Phonetics, University of Copenhagen 1 1967).

Martinet:

Comme le français semble, en matière d'assimilation de voix, en cours d'évolution, il serait bon de préciser l'âge des informateurs et également leur origine géographique car le nord et le sud, l'est et l'ouest, ne réagissent pas de la même façon en la matière.

Wittoch:

Il faut tenir compte aussi du *style*: par ex., dans un style *soutenu*, le mot „guide“ — dans les groupes de mots „guide savant“ et „guide zoulou“ — doit être prononcé, à notre avis, dans les deux cas d'une façon pareille, sinon exactement de la même façon.

ACOUSTIC FEATURES OF SYLLABILITY IN DANISH

JØRGEN RISCHÉL*

It is generally considered to be characteristic of the Danish language that it is not very distinctly articulated. This is undoubtedly true of unstressed syllables. The shwa vowel¹ of these syllables often fuses with surrounding voiced segments, and frictionless continuants often go together with an adjacent vowel to form long vowel-like sequences. Typical instances are forms like [hv:və], [y:ðə], and even trisyllabic forms like [hv:vəvə], [y:ðəðə]. When adjacent to a nasal or [l] the shwa segment is frequently dropped, especially medially, as in [fal(ə)nə], [bon(ə)nə], although such trisyllabic forms remain—at least in some kinds of Danish—distinct from dissyllabic forms like [falnə], [bonnə]. It is generally said that an adjacent consonant becomes syllabic when shwa is dropped (or fused) in such forms as those just quoted. However, it is not altogether clear how the distinction of syllable number actually manifests itself.

The concept of “syllability” being a crux in Danish phonetics it seems worthwhile to present the problem in its acoustic aspects. The findings surveyed in this paper are based on analyses of mingograph and spectrograph records.²

The core of the material used for this study is a series of 11 word pairs with an alleged difference of syllable number. The words were read aloud with sentence main stress in a standardized type of context in which they were not influenced by utterance-final intonation (the manifestations in other contexts often deviate quite much but are not dealt with here). Individual, meaningful sentences ([vɔ:ðə], [vɔ:əðə] in *jeg vil rode grundigt; og han roede dårligt*, etc.) as well as inclusion of the words in a standard carrier phrase has been used, the former more extensively.

For consideration of space, numerical data (in all cases average values) will be given only to exemplify the phenomena dealt with (a fuller treatment with documentation will be given elsewhere).

* Institute of Phonetics, University of Copenhagen.

¹ No distinction is made here between “vowel, consonant” and “vowel, contoid”.

² The speech of six subjects speaking a Standard Danish typical of Eastern Denmark has been analysed. The following statements are essentially based on data from four of these, the remaining two being too deviating in speed of utterance and other features to be directly comparable to the rest.

An obvious characteristic of consonants taking over the syllabicity function of shwa is that they are longer than nonsyllabic consonants. This has been pointed out by Otto Jespersen³ and more recently by Poul Andersen.⁴ According to my data the difference in duration is clear (and statistically significant) in some word types at least; however it is sometimes difficult to determine what part of the word is longer. For example, intervocalic [ln] in [fal(ə)nə] is longer than is [ln] in [falnə], the increment being typically of the order of 40–50 ms, but this extra length does not seem to be consistently associated with either [l] or [n] (although with two subjects at least it is rather in the first consonant). In fused sequences involving [ð, ʁ] it is difficult to give a realistic measure of segment duration, but the duration of the whole voiced portion (including the stressed vowel) has been measured. Trisyllabic forms are generally somewhat longer than corresponding dissyllabic ones,⁵ but the variation is considerable, and the overall duration can hardly be the only essential difference between them (with some words it is not even clearly significant).

Eli Fischer-Jørgensen⁶ has suggested that there is in [bon(ə)nə] (which she transcribes [bonnə]), an increase and decrease of intensity or “strength”, and that this—possibly together with longer duration—accounts for the impression of an extra syllable peak. In my material trisyllabic forms sometimes do exhibit a steep intensity peak more or less coincident with a peak in the fundamental frequency contour. The intensity peak may, however, be a feature of very distinct pronunciation;⁷ it does not seem to occur very constantly (in fused sequences particularly with [ð] there is no such distinction at all in the majority of cases). The distribution of this feature in respect to subjects and word types remains to be studied in detail.

Leaving aside differences of segment quality (which do occur, although they need not be there)⁸ we may turn to pitch as a possible cue to syllabicity. In the type of material examined stressed words generally have a somewhat rising pitch with a short (normally quite short) fall at the end of the word.⁹ Thus the peak of the pitch is generally close to the end of the word. However, trisyllabic words seem to differ from dissyllabic ones by having the peak normally much earlier, so that typically about two thirds of the word (excluding initial consonants) are occupied by the rise, and the last third by the fall. Typical examples of the difference are: [hv:ʁə] with

³ Modersmålets Fonetik (3rd ed. 1934), p. 115.

⁴ Dansk Fonetik (in: Nordisk Lærebog for Talepædagoger, 1954), p. 323.

⁵ An extra complication is the distinction of forms like [sy:əðə] vs. [sy:ðəðə], which is not accounted for in this paper.

⁶ Almen Fonetik (3rd ed. 1962), p. 94.

⁷ Cp. the distinction mentioned in note 5 above.

⁸ Also cp. note 5.

⁹ Some speakers of Danish do not seem to have this rise-fall but either a pure rise (this applies to one subject not included here) or a fall. For these, of course, there can be no difference of timing of the pitch peak. The general shape of the word contour is an important feature of regional pronunciations of Danish.

5,2 per cent of the duration taken up by the fall versus [hv:ʁəʁə] (fused) with 30,2 per cent, or [ʁo:ðə] with 9,6 per cent versus [ʁo:əðə] with 29,3 per cent (data from two different subjects, all average values). In such cases the difference is clearly significant. In a case like [bonnə] versus [bon(ə)nə] the same difference can be found, but sequences like these involving [n] or [l] present special problems which cannot be approached in the present paper.¹⁰

Fundamental frequency curves of Danish words have been recorded earlier by Ekblom, Sv. Smith and others, but the role of pitch in the distinction of syllable number has not been in the focus of attention.

A number of different cues seem to be present to a different degree under different conditions. In order to see if pitch inflection can at all serve as a useful cue to word structure I made a brief experiment. A small series of synthesized stimuli consisting of [h] plus [v:] and differing only in pitch contour (roughly imitating contours typical of [hv:ʁə] and [hv:ʁəʁə]) were listened to by some of my colleagues and myself, and each was identified as either an imitation of [hv:ʁə] or of [hv:ʁəʁə]. Responses were unanimous and as expected, i.e. pitch seems to somehow enter into the signaling of syllabicity in Danish, or rather in the variety of Danish dealt with here.

DISCUSSION

Fischer — Jørgensen:

As I was quoted for having said (in my Textbook of Phonetics) that syllabicity in the cases mentioned should be due to intensity, I just want to emphasize that this was not based on instrumental research and that I find Mrs. Rischel's results very convincing.

Wütting:

My question is whether it would not be more relevant linguistically to speak about morphemes and morpheme boundaries (some sort of internal juncture) rather than syllables (syllabicity) and syllable boundaries when dealing with cases such as Danish *faldne* versus *faldende*.

Smith:

In case your description of a typical intonation curve in your examples holds true, we may foresee a development of our Danish language into a tone language. This development would then be parallel to the development of the Stød in Old Danish, developing as a dynamic accent under similar conditions.

¹⁰ With some subjects there are very pronounced ups and downs of the fundamental frequency contour at segment boundaries, so that it is not always possible to use timing of the peak as a parameter. With one subject I got a clear difference in words with [n] and [l], the percentages being comparable to those for other word types (with [ð], [ʁ]), but with some of the others the curves are quite puzzling.

Would you please tell me what is the smallest difference in fundamental frequency in the synthetic words between the two intonation levels required to distinguish the minimal pairs?

Rischel:

Ad Frøkjaer—Jensen: I have not made perception tests aiming at the determination of discrimination thresholds. Such tests should obviously be carried out both for pitch, duration and possibly intensity. The informal test I mentioned only showed that fundamental frequency contours imitating those found in real speech (in cases that are clearly different), i.e. with a considerable fall phase for trisyllabic but not for dissyllabic words, *can* provoke a listener to identify the stimulus as different words [in this case *hårde* 'hard' (pl.) versus *hårdere* 'harder'].

Ad Jørgensen, H. P.: I am studying the realization of the word types in connected speech, but the data are complicated to deal with. The difference of pitch timing tends to disappear in positions after the word carrying the main sentence stress. Other factors as well influence the pitch contour quite considerably, such as: following unstressed syllables or absolutely final position. Duration may be more important in such cases.

Ad Witting: My paper was concerned with the problem of what the phenomenon which we commonly refer to as syllabicity of consonants in Danish looks like from the acoustic point of view. However, I entirely agree that there is in the word pairs I have examined a difference of morphemic structure, and that the signalling function of pitch and duration is associated with this structural distinction.

INTONATION AND PREDICTABILITY

ALEXANDRA ROCERIC—ALEXANDRESCU

All linguistic statements consist of an initial pole—where we find the emitter—who produces the acoustic signal(s) and a final pole, of the receiver, who collects these signals, if transmitted to him under optimal conditions.

From the phonological point of view, we may divide and analyse the statements into two components: the phonic segments and the suprasegmental units. The intonation is a suprasegmental element: it has an universal character, because there do not exist languages in which the oral communication would be possible without intonation. By *intonation* we mean all the phonic elements which characterize the different realizations of the same phonematic chain.¹ “Non-intonational” characters have only those chains of phonemes considered “by themselves”, as abstract entities. By *prediction* we mean the selective process the listener effects upon receiving a statement; knowing the first part of the statement, he may “guess” more or less, the elements which will ensue, from an unknown message, the information supposing a choice between more possibilities (2).

We shall try to establish if the intonation may contribute or not to the “prediction” of the elements which compose a statement.

Referring to the intonational contours, we observe that, generally speaking, the final part of an intonation is marked by the characteristic tone of the final contour. In the Romanian language, for instance, we observe that the tone [2] characterizes the contour //, and [1] the contour /↓/. While the contour /↓/ may be or not accompanied by pause /#/, (it follows that the information will be obtained by the choice between two possibilities), in another case, between // and /#/ we have to note a relation of exclusion: where one of the two elements appears, the other is excluded (that means, the perception of the toneme supposes an exact information concerning

¹ Among these elements, we include first and foremost the melodic line (that is, the variations of the tonemes), the accent of intensity, the duration, the pause, the speech-rate, the rhythm, the intensity with which we utter the whole sentence, as well as other acoustic-physiological elements, which were not sufficiently analysed until now (for instance, the “tremolo”, characteristic to a certain theatrical speech); we include here also the concept of regional or foreign “accent”, as well as the variations of tone of the same sound in different pronunciations, a. s. o. All these elements are quoted by us as being facts of intonation (1).

the pause, too; the absence of pause indicates that the statement continues (3). We may multiply the examples. We mention, for instance, that in a dialogue, after an interrogation (or a series of interrogations) an enunciative sentence is to be expected, as an answer to the questions. The distinctions can be continued: we can "predict" the structure of the answer, which will be different when we have to do with a general question or with a *yes* or *no* question. Some research-workers have also pointed out the existence of intonations characterizing different types of sentences (4). When we follow an enumeration, we "recognize" a specific intonation of enumeration and we are able to "predict" the elements which will follow (in a more accurate approximation, we have the intuition of homogene sequences, as: nouns, adjectives, etc.—which belong to the same class).

A characteristic organization of the tonemes is to be noted for the disjunctive propositions, or for the conclusive ones, in Spanish, for instance. On stating this, we do not consider that a necessary bound does exist between the intonational contours and the different grammatical categories. The intonation of the first part of the statement "suggests" only the melodic units which are to complete the contour. By these means, also the elliptic constructions may be understood. But for the entire knowledge of the statement, other elements are important, as well.²

We must keep in mind that, if in the transmission of the information, the unpredictable character of some sequences which compose the texts may modify the intonational contours, there exists an inverse phenomenon also: the intonation may contribute to the increasing of the quantity of information, by means of increasing the predictability of the sequences with which the intonational contours are associated.

BIBLIOGRAPHY

1. Eisler—Goldman Frieda, A Comparative Study of Two Hesitation Phenomena, *Language and Speech*, IV, 1, 1961 (5).
2. Miller G. A., *Language and Communication*, New York—Toronto—London, 1959 (2).
3. Roceric—Alexandrescu Alexandra and Copceag D., *Sugestii pentru cercetarea structurală a intonației*, SCL 3/1966 (1).
4. Navarro Tomas, *Manual de entonación española*, ed. II, New York, 1948 (4).
5. Uldall Elizabeth, Attitudinal Meanings Conveyed by Intonation Contours, *Language and Speech*, III, 4, 1960 (6).
6. Vasiliu Em., *Fonologia limbii române*, Bucharest, 1965 (3).

² The most important is, surely, the context. We have to note the degree to which we know a certain language. Sometimes, intonation is associated with mimes, gestures, etc.—a series of communication components which we cannot discuss here and now. We shall also omit in our report the affective elements which come to the fore only by means of intonation and whose reception has already been described by some authors (5), (6).

A PHONETIC REALIZATION OF THE PAST MORPHEME IN ENGLISH

E. RODÓN

The importance of spoken language as basic material for linguistic study needs no longer to be emphasized. It is also well known that in the phonetic continuum of spoken language words are pronounced more or less linked to each other and that the final sound or sounds of any word may be altered by the initial sound of the following word. This phenomenon was perfectly well known by Sanskrit grammarians but has not always been taken into account in modern studies. Another fact we have to consider is that communication is the aim of linguistic activity. Consequently the reaction of listeners will be of the greatest importance in the discrimination of linguistic categories and the perception of the acoustic speech-signal has become one of the main points of interest in phonetic studies. In morphology we group into one morpheme not only phonemically identical but also phonemically different segments, provided they occur in different environments or if they do occur in the same environment they are free variants of each other. According to these criteria the past morpheme in English is said to have three allomorphs [-əd], [-d], [-t] which are phonologically conditioned by the following distributional pattern: [-əd] after roots ending in [-t, -d]; [-d] after roots ending in a voiced sound, except, [-d]; and [-t] after roots ending in [-p, -k, -f, -ð, -č, -š].

If speech means connected speech and we cannot take words in isolation, the environment of any final sound of a word is not only the preceding sound or sounds but the initial sounds of the following word as well.

Thus if a past tense of a verb has either of the two allomorphs [-t, -d] occurring after a plosive, a three consonant cluster will be produced in speech whenever the following word begins with another plosive. It is well known that in a two consonant cluster we do not pronounce both fully; we have implosion of the first consonant and the release of the second one; and that in a three consonant cluster there is the implosion of the first consonant, then a lengthening and finally the release of the third one. Thus a prosodic feature, length, has been substituted for the medial consonant. In our case, the medial consonant [-t, -d] happened to be the morpheme of past tense. Consequently in this particular environment, namely, after roots ending with a plosive, except [-t, -d], and followed by a word beginning with a plosive, the phonetic realization of the past morpheme in English is the prosodic feature of length.

There is still another fact to consider. In those three consonant clusters it can also

happen that, in very rapid speech, the medial consonant is not pronounced at all. In which case there would be no difference between the present and past tenses and we would have neutralization of the past morpheme. So the next thing to do is to find out what actually happens in current standard speech.

We made a list of pairs of sentences: one containing a past tense of the kind seen above (They walked twenty miles, The cats looked tame) alternated with the corresponding present forms. In the sentences, we tried to avoid any contextual guidance for the identification of tense, and in addition we gave the list to five different persons and asked them to cross out those sentences where they felt there was some indication of time in the context. And we selected the twenty pairs of sentences that had not been crossed out by any of the individuals.

Then we had these sentences recorded by a native English speaker unaware of what was involved in the experiment. In order to get a more lax pronunciation we gave him the sentences mixed with some other irrelevant material, and later on we re-recorded them again separately.

We took then two groups of native English listeners and made two different listening tests. We had the twenty pairs of sentences read to one of the groups; to the other group we played the recorded sentences. In both cases we asked the listeners to identify the present or past tense of the utterance. The evaluation of the results proved that in 12 sentences in the first group (read sentences) and in 9 in the second group (recorded sentences), the prosodic element present in the utterance had served as a signal of past tense, and that in the other cases (the opposition present) past tense had been neutralized in spoken language.

The grammatical implications of a prosodic feature being used as a morpheme have already been discussed by me in another paper of somewhat wider scope, here we are only concerned with the morphophonemic status of the element.

And the number of cases in which the prosodic feature of length was used as an effective means of signaling past tense in the communicative process entitles us to take it as an allomorph of the past morpheme in the particular environment we have described above.

VIBRANTEN ALS PHONEMKATEGORIE

MILAN ROMPORTL*

Gleich am Anfang muß betont werden, daß in diesem Referat die R-Laute in jenen Sprachen außer acht gelassen werden, wo sie unbestreitbar in allen Positionen als Spiranten (bzw. kontinuierliche Konsonanten) realisiert werden. Man muß sie dann innerhalb der Kategorie der Spiranten klassifizieren. Von anderen („üblichen“) Spiranten werden sie durch verschiedene andere akustische Merkmale als durch das „Flattern“ bzw. Schwanken der Intensität, Tonhöhe oder Klangfarbe unterschieden.

Die R-Laute in den übrigen Sprachen (z. B. in den slawischen oder den meisten romanischen, z. T. auch im Deutschen usw.) wurden als flatternde, gerollte u. ä. Konsonanten beschrieben und klassifiziert, solange man bloß artikulatorische Kriterien zur Unterlage der Klassifizierung benutzte. Der Übergang zu einem überwiegend oder ausschließlich auf akustische Kriterien gestützten Klassifikationsystem hat die Selbständigkeit der Kategorie der Vibranten in Zweifel gestellt. Die üblichen Methoden der akustischen Analyse, v. a. die am häufigsten benutzte sonographische Analyse, haben sehr oft keine eindeutigen Spuren eines den artikulatorischen Bewegungen entsprechenden Schwankens im akustischen Signal registriert, obwohl die Perzeptionsteste bewiesen haben, daß die Realisierungen der R-Phoneme auch in solchen Fällen so gut wie eindeutig als Vertreter der R-Phoneme von den Realisierungen anderer Phoneme desselben Sprachsystems unterschieden werden. (Die Streuung der Fehlerbeurteilungen entspricht jener bei anderen Phonemkategorien.) Auf Grund dieser Tatsachen ist man zu der Schlußfolgerung gekommen, daß der Hauptzug (bzw. das Hauptmerkmal), durch welchen die R-Phoneme von den Phonemen anderer Phonemkategorien derselben Sprache unterschieden werden, nicht das „Flattern“ oder „Rollen“, das Schwanken der Eigenschaften des Schalles während der Realisierung eines R-Phonems ist, sondern andere (inhärenten) Qualitäten des akustischen Signals.

Wir haben v. a. die Vibranten des Tschechischen einer gründlicheren Analyse unterzogen, deren v. a. das tschechische /ř/ betreffende Resultate in unserer Abhandlung im Sammelband *Phonetica Pragensia* veröffentlicht worden sind.¹ Deshalb wollen wir

* Phonetisches Institut, Karlsuniversität zu Prag.

¹ M. Romportl, Ř und das tschechische Konsonantensystem, *Phonetica Pragensia*, AUC — *Philologica* 6, Praha 1967, S. 7—27.

an dieser Stelle nur die Hauptergebnisse dieser Untersuchungen kurz zusammenfassen.

Wie bekannt, hat das Tschechische zwei Vibranten, /r/ und /ř/, wobei das zweite Phonem in zwei positionsbedingten Varianten (Allophonen) vorkommt, dem stimmhaften [ř] und stimmlosen [ř°]. Die Realisierung dieses letzteren Phonems, die Artikulation und Akustik, weniger schon Perzeption seiner Varianten erregt seit Jahrzehnten die Aufmerksamkeit der Phonetiker und Phonologen, von denen wir aus den letzten Jahren v. a. die Arbeiten von Vachek,² Hála,³ Kučera,⁴ Zima⁵ und zuletzt Isačenko⁶ erwähnen. Isačenko's Aufsatz, wo er die Aussprache einer Versuchsperson sonographisch analysiert, hat überzeugend gezeigt, daß wenigstens in einigen Positionen bzw. bei einigen Sprechern das tschechische [ř] nicht als Flatterlaut ausgesprochen wird. Er ist zu dem Schluß gekommen, daß man [ř] als einen „affrizierten Sonorant mit dem Merkmal scharfklingend (strident)“ interpretieren sollte.

Indem unsere Analyse der Realisierungen des tschechischen /ř/ gezeigt hat, daß die Schwankung der Qualitäten des Schalles (Tons) — v. a. der Intensität, aber gewöhnlich auch des Grundtons sowie der Obertöne, bzw. der „Klangfarbe“ — regelmäßig vorhanden ist, die Analyse des [ř] und [ř°], die wir durchgeführt haben, hat unter Anwendung eines viel weiteren Materials die von Isačenko konstatierten Tatsachen nur in einigen Fällen bestätigen können. Sie hat z. B. gezeigt, daß die Realisierungen des /ř/ zum größten Teil eine Schwankung der einzelnen Qualitäten des Schalles aufweisen (z. B. des Grundtons im stimmhaften Allophon, des Schalldrucks, des akustischen Spektrums in beiden Varianten),⁷ sehr oft hat sie aber ergeben, daß das Schwanken nicht unbedingt alle Qualitäten gleichzeitig betrifft und — hier im Einverständnis mit Isačenko —, daß es manchmal nur in einem Teil der Realisierung oder überhaupt nicht vorkommt. Dabei war in einigen Fällen das akustische Spektrum sowie das Zielen des zweiten (bzw. auch der weiteren) Formanten der benachbarten Vokale (Lokus) von dem Spektrum und Lokus anderer Zischlaute verschieden.

Man könnte — wenn man unsere an mehreren Stellen früher formulierte Hypothese über die Struktur der distinktiven Merkmale einfach applizieren sollte⁸ — zu der Schlußfolgerung geraten, daß das Phänomen des Flatterns (bzw. Schwankens) in

² J. Vachek, The place of the sound [ř] in the structures of slavonic languages, *Sbor. prací filol. fak. brněnské univ.* 12, 1963, A 11, S. 81—92.

³ B. Hála, Sur la hiérarchie des éléments phoniques des sons du langage, *Proc. Münster*, Basel 1965, S. 332—335 (auch in *Slavica Pragensia* 4, Praha 1962, S. 63—68).

⁴ H. Kučera, *The phonology of Czech*, 's—Gravenhage 1961.

⁵ P. Zima, Souhláska Ř v českém systému znělostní asimilace, *Sborník slavistických prací*, Praha 1958, S. 36—43.

⁶ A. V. Isačenko, Zur Akustik des tschechischen ř-Lautes, *Phonetica* 12, 1965, S. 1—12.

⁷ Vgl. Abb. 1 bis 15 in unserem bereits erwähnten Aufsatz.

⁸ Vgl. M. Romportl, K metodice zvukového rozboru jazyka, *Slovo a slovesnost* 17, 1966, S. 208—216, bzw. auch Zu akustischen Korrelaten der distinktiven Merkmale, *Proc. Münster*, Basel 1965, S. 506—511, bzw. *Zvukový rozbor ruštiny*, Praha 1962, u. a.

der Hierarchie der ein distinktives Merkmal bildenden akustischen Züge sehr tief stehen muß.

Man darf aber die in der Prager Schule seit Mathesius so oft betonte Potenzialität der sprachlichen Erscheinungen⁹ nicht außer acht lassen. Unsere Hörteste haben gezeigt, daß z. B. ein im Zeitverlauf regelmäßig (zwei- bis fünfmal) gedämpftes Zischlautspektrum (nicht nur das Spektrum eines [ř] oder [ž], sondern auch [s] oder [z]) von dem tschechischen Hörer als entsprechende Variante des [ř] identifiziert wird, wenn gleichzeitig auch der Verlauf der benachbarten Transientphasen dem Verlauf vor oder nach [ř] bzw. [ř°] entspricht.

Dagegen wird eine Verlängerung eines Spektrums, das dem des [ř] oder [ř°] entspricht, ohne Dämpfungen nicht eindeutig wahrgenommen und identifiziert.

Zu ähnlichen Ergebnissen sind wir gekommen, als wir das L-Spektrum einmal bis dreimal durch eine Dämpfung unterbrochen haben, wobei der Hörer einen R-Laut perzipiert hat.

Diese abwechselnde Dämpfung (sowie Verminderung des Umfangs anderer akustischer Eigenschaften), die im natürlichen Laut nur potenziell ist, muß unserer Ansicht nach trotz seiner Potenzialität für den Hauptzug der Vibranten gehalten werden, da in ihrer Anwesenheit der Laut nur als Vibranta wahrgenommen werden muß. Sie wird selbstverständlich auch durch andere akustische Züge begleitet, die im Falle ihres Nichtvorhandenseins dafür supplieren können. Keiner dieser anderen Züge funktioniert aber so eindeutig.

⁹ V. Mathesius, *O potenciálnosti jevů jazykových*, Praha 1911.

SUR LA HIÉRARCHIE DES PARAMÈTRES DE L'ACCENT

MARIO ROSSI*

Parmi les phonéticiens qui se sont occupés de dégager le paramètre dominant qui rend compte du relief accentuel, on peut distinguer les partisans de l'intensité, les partisans de la durée et ceux de la hauteur. Pour le détail des noms des tenants de l'une ou l'autre de ces théories, nous renvoyons aux excellentes bibliographies de Bolinger et de Faure dans leurs ouvrages sur cette question.¹

La plupart des phonéticiens s'accordent aujourd'hui à dire que le facteur déterminant de la proéminence accentuelle est la hauteur. Ces chercheurs, aux prises avec une réalité complexe et souvent décevante à analyser, ont en général orienté leurs investigations vers la synthèse de la parole. Les résultats acquis dans ce domaine ne prouvent cependant pas que dans le langage réel la fréquence soit le seul „caractère physique pertinent“ ou même le paramètre dominant de l'accent. De plus, certains travaux confondent les deux niveaux intonatif et accentuel: on découvre parfois dans l'acte de parole, ou l'on met en évidence dans une phrase synthétique, un paramètre — la hauteur — que l'on attribue à l'accent mais qui n'est autre que la réalisation de l'intonème progrédient, terminal, ou même interrogatif.² Une autre erreur consiste à définir l'accent sur le seul plan perceptif comme une proéminence: on risque d'aboutir ainsi, en français, à une confusion linguistique entre l'accent d'expression et l'accent dit „rythmique.“ L'organisation prosodique de la phrase, dans les langues à accent libre, permet d'éviter la confusion entre les niveaux intonatif et accentuel: en italien, par exemple, le groupe identifié par l'intonème progrédient ou terminal, peut comporter plusieurs accents: un accent principal qui se réalise en même temps que l'intonème, un ou deux accents secondaires qui se réalisent en dehors de celui-ci. La langue sur laquelle porte cette étude, dialecte liguro-émilien parlé au nord de la province de Massa-Carrara, présente sur ce point la même structure que l'italien.

* Faculté des Lettres et Sciences humaines, Aix en Provence.

¹ Bolinger, D. L., A Theory of pitch accent in English, *Word*, 14, 1958, PP. 109—149, repris dans *Forms of English*, Tokyo, 1965, pp. 17—55. Faure, G., *Recherches sur les caractères et le rôle des éléments musicaux dans la prononciation anglaise*, Didier, Paris, 1962, pp. 125—151.

² Parmenter, C. E. and Blanc A. V., An experimental study of accent in French and English *PMLA*, XLVIII, pp. 598—607; voir aussi Rigault, A., La perception de l'accent en français, *Proceedings of the fourth international congress of phonetic sciences*, Helsinki, 1961, p. 735.

Nous nous attacherons par conséquent à définir les paramètres de l'accent intérieur qui se réalise soit sur la dernière, soit sur l'avant-dernière syllabe du mot. Nous verrons ensuite si nous retrouvons les mêmes paramètres dans la réalisation de l'accent principal; la comparaison entre accent principal et accent secondaire nous permettra de dégager les facteurs qui rendent compte du supplément de force de l'accent de fin de groupe et d'établir une hiérarchie des paramètres responsables de la prééminence de l'accent tel que nous venons de la définir. On ne saurait nous reprocher de limiter nos investigations au plan acoustique, voire physiologique, car nous nous plaçons sur le plan auditif, du fait que nous tenons compte de la valeur spécifique des sons³ et de la façon dont l'oreille intègre les unes en fonction des autres la durée, l'intensité et la hauteur.

a) Nous ajouterons à l'intensité de chaque voyelle le facteur de correction suivant:

a	â	ə	è	ã	é	eɪ	ö	o	ō	ø	ē	y	u	i
0	0	-1	+1	+1	+2	+2	-2	+2	+3	+3	+3	+3	+3	+5

b) Les intensités ainsi corrigées seront converties en phones en fonction de la valeur du fondamental.⁴

c) Cette réduction ne suffit pas à convertir l'intensité en sonie: reste le problème, à notre avis capital, de l'intégration temporelle de l'intensité. Faure et Bolinger ont fait allusion aux rapports entre durée et intensité.⁵ Les travaux des psychoacousticiens sur les rapports durée — intensité doivent être pris en considération.⁶ L'étude la plus appropriée à notre sujet est celle de Munson,⁷ qui a le mérite pour nous d'indiquer de façon précise la perte d'énergie en phones pour des durées qui sont celles des voyelles dans la conversation courante. Il donne les valeurs de l'intensité perçue pour un son de 1000 cycles à 70 db, en fonction de la durée depuis 0,25 S. jusqu'à 0,01 S.; entre 0,20 et 0,25 S. l'intensité du son est perçue pratiquement dans sa totalité;

³ House, A. S., and Fairbanks, G., The influence of consonant environment upon the secondary acoustical characteristics of vowels, *JASA*, 25. 1. 1953, pp. 105—113; Lehiste, I., and Peterson, G. E., Vowel amplitude and phonemic stress in american English, *JASA*, 31. 4. 1959, pp. 428—435; Ladefoged, P., Sub-glottal activity during speech. *Proceedings of the fourth international congress of phonetic sciences*, Helsinki, 1961, pp. 73—91.

⁴ cf. *Lignes isosoniques normales pour sons purs écoutés en champ libre et seuil d'audition binauriculaire en champ libre*. Norme française, S 30—003, septembre 1965, AFNOR, 23, rue N. D. des Victoires, Paris 2^e.

⁵ Faure, pp. cit., p. 143.

⁶ Pollack, I., Loudness of periodically interrupted white noise, *JASA*, 30, 1958, pp. 181—185; Garner, W. R., The effect of frequency spectrum on Temporal integration of energy in the ear, *JASA*, 19. 5. 1947, pp. 808—815, donne comme valeur critique au-delà de laquelle l'intensité perçue ne change plus 0,25 s. Voir également Mikolezy-Fodor, F., Relations between loudness and duration of tonal pulses, *JASA*, 31, 1959, pp. 1128—1134, et Small, A. M., Brandt Hr, J. F., and Cox, P. G., Loudness as a function of signal duration; *JASA*, 34, 1962, p. 513.

⁷ Munson, W. A., The growth of auditory sensation, *JASA*, 19. 4. 1947, pp. 584—591.

au-dessous de 0,10 S. la sonie décroît très rapidement (fig. 1). Si l'on sait que les voyelles inaccentuées ont une durée moyenne de 0,06 s., que les voyelles accentuées de fin de groupe dépassent 0,20 s., la simple lecture de ce tableau nous éclaire, à première vue, sur les raisons de la prééminence accentuelle: la différence de sonie entre les deux catégories de voyelles est en effet de 10 phones!

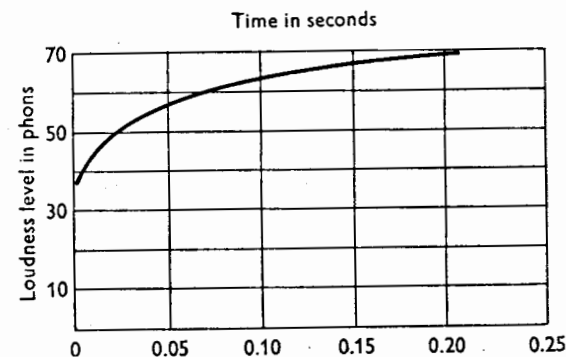


Fig. 1. Loudness level of a 1000 cycle tone as a function of duration time (intensity level = 70 db). (Extrait de Munson, p. 586, fig. 37).

A. *L'accent secondaire*.⁸ Dans le corpus étudié nous remarquons que la durée est, avec l'intensité, le paramètre le plus souvent présent, que si nous considérons l'accent principal elle est *toujours* présente. Le calcul de l'intensité subjective en fonction des corrections mentionnées (fig. 2) montre que la sonie de la voyelle sous l'accent secondaire est plus forte que celle des voyelles environnantes: dans 28 cas sur 29, la sonie de la voyelle accentuée est plus forte que celle de la voyelle précédente, d'au moins 2 phones. Dans 36 cas sur 40 elle est plus forte que celle de la voyelle suivante, d'au moins 2 phones. Lorsque la syllabe accentuée est précédée et suivie d'une voyelle, la sonie de son noyau sonore domine à la fois les voyelles environnantes dans 25 cas sur 29. La dominance moyenne de la sonie est très confortable puisqu'elle est de l'ordre de 8 phones (cf. fig. 3).

Par contre, nous n'observons de rupture mélodique sous l'accent que dans 49% des cas (19 sur 40). Nous pouvons donc affirmer, avec 1 chance sur 100 d'erreur que la sonie est le paramètre constamment présent qui rend compte du relief de l'accent

⁸ Sujet: homme, 45 ans, habitant Rossano (haute vallée à 15 km ouest de Pontremoli, prov. Massa-Carrara); enregistrement réalisé dans la chambre anéchoïque du CBRS à Marseille, sur magnétophone Nagra 3B, microAKG, D 24/200/B. La fréquence fondamentale est analysée par détecteur de mélodie, la courbe de niveau sonore fournie par intensimètre logarithmique (ces appareils ont été conçus et réalisés par MM. Lancia, Carré, Paillé et Beauviala de l'Ecole d'Ingénieurs électroniciens de Grenoble); les données sont reproduites sur papier à l'aide de l'oscilloscope Siemens. Notre corpus comprend 71 groupes, dont 45 groupes progressifs et 26 groupes terminaux; nous avons identifié au total 40 accents secondaires.

		n	d	ã	m	a	'p	i	a/l	a	m	ʔa	k	i	n	a
1	fréquence fondamentale	157	153	160	160	170	190	145	130							
2	durée en s.	0.04	0.04	0.06	0.06	0.06	0.20	0.08	0.08							
3	intensité en db	63	64	64	66	66	64	59	50							
4	intensité spécifique	0	0	+5	0	0	0	+5	0							
5	total	63	64	69	66	66	64	64	50							
6	intensité en phonies en fonction du fondamental	61	62	68	65	66	64	62	43							
7	intégration temporelle de l'intensité	-14	-14	-11	-11	-11	0	-8	-8							
8	total 2 = sonie	47	48	57	54	55	64	54	35							

Fig. 2. Accent secondaire sur [p i], accent principal sur [ma]. Les chiffres de la ligne 3 représentent les valeurs de l'intensité au-dessus du seuil 10—16 watts/cm², l'enregistrement ayant été réalisé avec une calibration en niveau (1000 à 65 db, à 30 cm du micro). On remarque le relief bilatéral de la sonie sur chaque voyelle accentuée. Les valeurs du fondamental, par contre, sont intégrées à une courbe qui monte progressivement vers l'accent principal; la hauteur ne peut donc rendre compte de la prééminence de l'accent secondaire.

	Accent secondaire	
	Rapport en phonies entre V. accent. et V. précédente	Rapport en phonies entre V. accent. et V. subséquente
Moyenne et limite de confiance	8,57 ± 2,26	7,37 ± 2,41
N	26	37
Seuil	0,02	0,02

Fig. 3.

secondaire. Dans 92 % des cas, la sonie augmente progressivement jusqu'à la voyelle accentuée; dans la syllabe subséquente, elle diminue de quelques phonies puis augmente de nouveau jusqu'à l'accent suivant, comme le montre la figure 4.

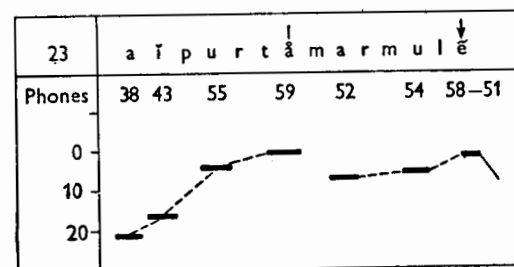


Fig. 4. Le symbole [↓] indique l'accent secondaire; [↓] l'intonème terminal; [↑] l'intonème progressif.

S'il est vrai, comme l'affirme en particulier Ladefoged, que l'auditeur est sensible à l'effort physiologique fourni par le locuteur, que l'intensité perçue est proportionnelle à cet effort, nous devons admettre que l'accent est bien, pour employer la formule de Stetson, le résultat de la culmination d'un mouvement. La sonie telle que nous la calculons, en effet, peut être considérée comme l'image de l'effort physiologique fourni par le locuteur. Les syllabes accentuables qui suivent la progression de ce mouvement sont désaccentuées; elles sont proéminentes lorsqu'elles se détachent de la courbe accentuelle grâce à une sonie à dominance bilatérale. Ce phénomène explique le fait que deux syllabes contigües ne puissent être accentuées à l'intérieur du même groupe, il explique également la perception d'une proéminence sur la première syllabe du groupe, lors d'une attaque forte.

Si la fonction accentuelle est remplie par la courbe de sonie qui est elle-même le reflet d'un mouvement physiologique, on peut se demander dans quel rapport elle se trouve avec la ligne tonale. On remarque que dans 50 % des cas, dans notre corpus du moins, l'accent n'est pas accompagné de rupture tonale. En général, la hauteur de chaque voyelle s'intègre à une courbe qui progresse vers la syllabe oxytonique ou paroxytonique du groupe, sur laquelle elle forme le signifiant de l'intonème progressif ou terminal. La ligne tonale joue donc essentiellement un rôle au niveau intonatif. On comprend dès lors pourquoi souvent la courbe de sonie et la courbe tonale présentent des schémas inversés.

B. L'accent principal.

1. Les groupes progressifs: La sonie de la voyelle accentuée calculée bilatéralement par rapport au contexte immédiat, est représentée dans la fig. 5. La sonie de la voyelle accentuée de fin de groupe est toujours plus forte que celle des voyelles environnantes, d'au moins 4 phonies. L'intensité subjective moyenne dans la réalisation de l'accent principal est bien supérieure à celle de l'accent secondaire. On

	Accent principal: groupes progrédients	
	Rapport en phones entre V. accent, et V. précédente	Rapport en phones entre V. accentuée et V. subséquente
Moyenne et limite de confiance	12,97 ± 2,26	16,51 ± 3,79
N	45	32
Seuil	0,02	0,02

Fig. 5.

pourrait penser que la sonie est seule responsable du supplément de force de l'accent de fin de groupe.

2. Groupes terminaux: La sonie de l'accent principal se situe en moyenne à 6 phones au-dessus de celle de la voyelle précédente et à 30 phones au-dessus de la voyelle subséquente (fig. 6). Le rapport de 30 phones entre les deux dernières voyelles du

	Accent principal: groupes terminaux	
	Rapport en phones entre V. accent, et V. précédente	Rapport en phones entre V. accent, et V. subséquente
Moyenne et limite de confiance	6 ± 2,51	30 ± 9
N	26	15
Seuil	0,02	0,02

Fig. 6.

groupe terminal, bien qu'étant très supérieur à celui que nous trouvons entre les mêmes voyelles du groupe progrédient (16,51 phones), ne semble pas apporter d'information supplémentaire sur le plan accentuel, puisque l'accent principal est moins proéminent dans le groupe terminal que dans le groupe progrédient. Par contre, il semble bien que la valeur du rapport entre la voyelle accentuée et la voyelle précédente rende compte en partie de la différence de proéminence entre les accents

principaux des groupes terminal et progrédient. Mais le relief de l'accent principal du groupe terminal est toujours plus grand que celui de l'accent interne. Or:

— le rapport de sonie accent principal /V. préaccentuée (6 phones) est, dans le groupe terminal, inférieur au rapport accent secondaire /V. (préaccentuée) $8,57 \pm 2,26$, au seuil 0,02.

— d'autre part, dans 50 % des cas (9/18) le rapport de sonie accent principal/accent interne s'établit en faveur de l'accent interne: la sonie calculée sur l'accent principal est inférieure à celle de l'accent interne. Dans les autres cas le rapport de sonie accent principal/accent interne n'atteint en moyenne que 5,44 phones.

Enfin la montée mélodique ne joue aucun rôle dans le groupe terminal où la hauteur de la voyelle sous l'accent principal se situe généralement au niveau de la ligne tonale de base.

Il ne reste par conséquent qu'un élément qui puisse rendre compte de la différence de proéminence entre l'accent principal et l'accent secondaire; cet élément qui est toujours présent sur la voyelle accentuée de la fin du groupe est le signifiant de l'intonème:

— le ton statique, dans le groupe progrédient

— la chute dont la valeur se situe entre la quarte et la quinte, sur la voyelle accentuée, dans le groupe terminal. La forme de la ligne tonale à la fin du groupe a essentiellement une fonction intonative. Mais en vertu de son association constante avec le dernier accent du groupe et à cause du haut degré de perceptibilité du facteur mélodique, l'intonème confère à la syllabe sur laquelle il se réalise une proéminence que ne peut avoir l'accent secondaire.

Nous pouvons dire, pour conclure, que l'accent qui permet de définir ce qu'on appelle l'unité accentuelle (unité qui ne coïncide pas toujours avec le mot) est, sur le plan physiologique et sur le plan auditif, le résultat de la culmination d'un mouvement: la sonie — image de l'effort physiologique fourni par le locuteur, est le facteur constant qui rend compte du mécanisme de la perception accentuelle; une rupture mélodique peut accompagner la sonie et faciliter dans une large mesure la perception du relief; au niveau du groupe, l'intonème est responsable de la distinction entre accent interne et accent final.

Le signifiant de l'intonème joue donc également un rôle au niveau accentuel: c'est le facteur dominant qui permet de distinguer l'accent principal de l'accent secondaire; ce qui revient à dire que c'est l'intonème et non l'accent qui a une fonction démarcative au niveau du groupe, que c'est l'intonème et non l'accent qui permet d'opposer l'unité accentuelle au groupe, tandis que l'accent est l'intégrant de l'unité accentuelle.

DISCUSSION

Carton:

On demande à M. Rossi de préciser comment il a procédé pour calculer le coefficient de correction des voyelles de son parler en vue des calculs de l'intensité.

Delattre:

Pourriez-vous préciser la différence que vous faites entre l'accent secondaire et l'accent principal dans le dialecte que vous étudiez? J'aimerais aussi savoir jusqu'à quel point ce que vous avez trouvé peut s'appliquer au français.

Gsell:

Je félicite Monsieur Rossi pour sa communication si riche et si précise. J'ai été très sensible au fait qu'il soit parti de la parole naturelle et qu'il en ait objectivement analysé les facteurs. On a trop tendance à croire que dans toutes les langues la mélodie est un facteur de discrimination de l'accent. J'aurais simplement voulu demander à Monsieur Rossi, pour illustrer son exposé, de nous donner rapidement un exemple de délimitation de syntagme par l'intonème dans son parler.

Landercy:

Les syllabes accentuées sont caractérisées par une augmentation relative de la sonie. Une telle augmentation peut provenir de la variation de trois paramètres — la durée, l'intensité, la fréquence. L'action d'un de ces trois paramètres est-elle prédominante?

Wajskop:

Les résultats de Munson concernent avant tout les sons purs. Est-il justifié d'appliquer cette pondération à la sonie de sons complexes? Pouvez-vous reprendre les chiffres obtenus par Garner en ce qui concerne la relation durée — intensité?

Rossi:

Ad Carton: La correction apportée pour rendre compte de l'intensité spécifique est une approximation de la réalité. J'ai extrait de plusieurs corpus des voyelles de même durée et de même hauteur (à 5/2 hz près) provenant de contextes identiques. Les sujets testés ajustent la sonie de chaque voyelle sur celle de [a]. La différence moyenne d'intensité entre [a] et les autres voyelles constitue le facteur de correction.

Ad Delattre: L'accent principal est l'accent fort qui se réalise à la fin du groupe. L'accent secondaire est un accent plus faible qui se réalise à l'intérieur du groupe, en dehors de l'intonème. La distinction entre accent secondaire et accent principal vaut pour les langues à accent libre. Je ne pense pas qu'elle puisse valoir pour le français; en français, l'accent phonétique semble bien se réaliser toujours en même temps que les intonèmes terminal ou progrédient. Je pense cependant que mon hypothèse sur le rôle de la sonie dans la mécanique accentuelle vaut également pour le français.

Ad Gsell: Voici un exemple simple qui permet de préciser les rôles respectifs de l'accent secondaire et de l'intonème [sə'koj/al rize] [sə'koj] forme une unité accentuelle ou syntagme; l'élément démarcatif est l'accent dont la prééminence est due à la sonie; sur [rize] l'intonème progrédient [↑], est l'élément démarcatif qui permet d'identifier le groupe.

Ad Landercy: Le paramètre le plus souvent présent est

1. la durée dans 75% des cas,
2. l'intensité dans 67% des cas,
3. la hauteur dans 50% des cas environ.

Ad Wajskop: Les résultats de Munson sont confirmés pour les sons complexes par les travaux de Pollack (voir bibliographie); nous pouvons donc penser que l'intégration de l'intensité se fait de la même façon pour les voyelles; il reste à le prouver; mais déjà les résultats auxquels j'aboutis semblent confirmer mon hypothèse.

Garner donne la durée 0.25 s. comme valeur limite au delà de laquelle la sonie n'augmente plus.

LA DESCRIPTION PHONÉTIQUE DE LA NÉGATION FRANÇAISE ET LA STYLISTIQUE

J. ŠABRŠULA

La réduction du *signifiant discontinu* de la négation française ne saurait être expliquée d'une manière trop mécaniste. On ne s'étonne pas de voir s'affaiblir et même disparaître le segment non accentué par lequel la négation s'exprimait en latin; c'est la constitution des groupes rythmiques dans lesquels la dernière syllabe de chaque groupe est forte, accentuée, qui explique en partie d'abord l'origine du caractère obligatoire d'une double négation dans une langue qui a hérité du latin la négation placée avant le verbe et qui a changé la place de l'accent, mais la réduction du signifiant discontinu de la négation française s'explique aussi par *contagion sémantique* (la première particule ayant passé la valeur négative à la seconde) et, ensuite, par le fait que l'élément *ne*, attaché à la forme du verbe prédicatif, est devenu particule d'introduction et *redondante* qui marque l'accord du verbe avec la particule négative principale. Mais, avant tout, c'est le caractère rhématique du second élément de la négation, qui a permis un écrasement phonétique de la partie redondante du signifiant discontinu de la négation dans le langage populaire et familier.

L'écrasement phonétique de la partie *ne* est le résultat d'une succession d'allophones entre [n] articulé et zéro. Du point de vue articulatoire et psychologique de l'émetteur, l'étape de l'écrasement phonétique de *ne* devrait être atteinte plus tard que du point de vue de la phonologie acoustique, où seules l'absence et la présence de voix perceptible peuvent être pertinentes (et non pas une infinité de degrés de chuchotements, quelquefois imperceptibles, jusqu'à la voix). Pratiquement toutefois, celui qui parle a également une expérience d'auditeur, ce qui facilite la synchronisation de la progression du phénomène chez l'émetteur et l'auditeur.

L'écrasement phonétique de [n] peut entraîner l'apparition de quelques allophones combinatoires dans l'entourage.

La négation contradictoire peut être exprimée aussi lexicalement. Souvent, elle est exprimée au moyen de certains éléments formateurs synthétiques spécialisés plus ou moins.

La plupart des couples à valeur contradictoire ou contraire représentent alors des oppositions binaires, formés par un terme de base, sans aucune marque spéciale, et un terme marqué formellement. Il s'agit d'oppositions privatives (*offensif*, *inoffensif*). Certains des antonymes ainsi formés n'ont pas toujours les mêmes valeurs

stylistiques (*quiétude, inquiétude*). Certains antonymes, considérés synchroniquement, présentent, outre une marque préfixée, une altération phonétique introflexive (*facile : difficile*).

Dans d'autres cas cependant, le schéma binaire ne suffit pas (*arranger : déranger : ranger*), quelquefois on trouve des oppositions équipollentes (*emballer : déballer*), dans certains cas il y a plusieurs synonymes, dérivés ou périphrases, en opposition avec un terme de base, etc.

Dans certains cas, l'existence ou la non-existence du terme négatif de base dérange la symétrie d'une famille de mots apparentés étymologiquement et unis par les affinités de forme et de sens (je pense à l'existence de certains termes négatifs qui n'ont pas de contraire affirmatif).

Les formes négatives se distinguent — dans la majorité écrasante des cas — des formes positives par l'addition d'un élément matériel, alors qu'il n'existe aucun cas où cette relation soit inversée. Ceci nous rappelle la formation du pluriel dans différentes langues, qui se distingue souvent du singulier par l'addition d'un morphème, alors qu'il n'existe aucune langue où cette relation soit inversée.

S'agit-il ou ne s'agit-il pas ici de motivation diagrammatique (c'est-à-dire d'icônes opérant par similitude de fait entre signifiant et signifié) une similitude de relation quantitative, le signifiant du pluriel p. ex. tendant à répondre à la signification d'une augmentation numérique par un accroissement de la longueur de la forme (Jakobson, Peirce)?

En ce qui concerne la négation, l'analogie consiste dans le fait que le terme nié renferme, du point de vue du contenu, le terme positif et apporte, en plus, une information supplémentaire (négation, contradiction du terme donné).

On pourrait objecter que cette observation pourrait être appliquée à la relation de n'importe quel terme non marqué et marqué d'une opposition linguistique quelconque.

De plus, les formes matérielles du pluriel diffèrent d'une langue à l'autre, ainsi que les formes de la négation. Alors il ne s'agit que d'un cas spécifique de motivation secondaire, analogique, déterminée par le système et l'existence des moyens sonors spécifiques de la langue donnée. On trouve aussi des cas inverses, où l'augmentation ou la diminution matérielle ou l'appauvrissement et l'enrichissement notionnel du contenu désigné sont en relation inversée avec la longueur de la forme du signifiant.

En résumé, disons que l'expression linguistique, quoique non motivée directement (*physei*), dépend du contenu, de la dynamique de l'énoncé, de la perspective fonctionnelle de l'énoncé. Les éléments phoniques au service du rhème sont pertinents. D'autre part, l'écrasement phonétique d'un élément phonique n'est pas en relation directe avec la disparition d'une fonction. Celle-ci peut persister, en chargeant un autre élément phonique. La réalité ou l'écrasement d'un élément phonique n'aboutit donc pas nécessairement à mettre en cause la primauté de l'ensemble.

Une dernière remarque: Dans une langue où le rhème de l'énoncé peut être

marqué aussi par l'intonème, ou bien dans laquelle le rhème attire certains intonèmes, c'est l'absence ou la présence de celui-ci qui peut — indirectement — entraîner la conservation ou la perte d'une partie du matériel phonique; la perte, surtout dans le cas d'une redondance évidente.

L'expression et le contenu ne sont pas mutuellement indépendants, dans certains cas il est possible de prouver une dépendance indirecte qui va du contenu, de sa substance et de sa forme, à la forme et la substance de l'expression.

DISCUSSION

Wittoch:

Dans la langue familière, on ne dit pas en français „je ne sais pas“ ou „je sais pas“, mais [*ʃepa*]. Le mot „je“ n'existe plus ici que comme une sorte de „trait distinctif“, par opposition à [*ty sɛ*], [*il sɛ*] etc.; autrement, ce mot a disparu.

L. Novák:

La condition structurale préalable de la création d'une double négation en français était, à mon avis, l'existence des doubles éléments morphologiques exprimant les personnes dans le système du verbe (quelquefois conçus comme morphèmes brisés): 1° en tant qu'éléments non-marqués les anciens pronoms *je... nous...*; 2° en tant qu'éléments marqués des désinences personnelles —*σ... —ō...* L'élément non-marqué de la négation *ne(n')* coïncide, selon sa position et dans le cas fondamental sans inversion, avec les anciens pronoms, tandis que son élément marqué *pas, point, guère, jamais...* coïncide de la même façon avec les désinences personnelles. La réduction secondaire de la négation, dans l'usage populaire, surtout parisien, supprime, dans le cas d'expressivité et même ailleurs, l'élément non-marqué (le type *sɛ pá* au lieu de *žə n[ə] sɛ pá = je ne sais pas*).

ФОНОЛОГИЧЕСКАЯ СИСТЕМА КАК ЕДИНСТВО КОМПЛЕМЕНТАРНЫХ ДИАКРИТИЧЕСКИХ СИСТЕМ

К. С. ШАУМЯН*

В истории теоретической фонологии существенное значение имели две фонологические теории: 1. реляционно-физическая теория (пражская школа), трактующая фонему и другие фонологические единицы как форму в физической субстанции; 2. глоссематическая теория фонемы, трактующая фонему и другие фонологические единицы как чистую форму.

Обе эти фонологические теории наталкиваются на фундаментальные трудности: первая из них ведет к открытой мною антиномии транспозиции; вторая игнорирует взаимозависимость формы и субстанции, поскольку к форме фонологических единиц безусловно относятся дифференциальные элементы и их парадигматические и синтагматические характеристики, находящиеся в прямой зависимости от физической природы звуков и, в частности, от линейных свойств физического речевого потока.

Чтобы преодолеть трудности, с которыми сталкиваются обе теории, мною был выдвинут принцип двухступенчатости. Согласно этому принципу единицы фонологической системы должны быть расщеплены на комплементарные диакритические элементы: фонема и фонемоиды, дифференторы и дифферентоиды, просодемы и просодемоиды и т. д. В результате мы получаем двухступенчатую систему, состоящую из комплементарных диакритических систем: абстрактной диакритической системы и реализующей ее функциональной диакритической системы.

Принцип двухступенчатости — это методологический постулат, отличный от различных уровней абстракции в работах некоторых фонологов (ср. например, различие эпического и этического уровней у К. Пайка и т. п.). Существенным для двухступенчатой теории фонологии является то, что здесь фонемы рассматриваются не как инварианты звуков, а как единицы более высокого порядка. Обе диакритические фонологические системы имеют свои собственные инварианты и варианты. Поэтому принцип двухступенчатости называется мною также принципом гетероинвариантности.

Для исследовательской практики двухступенчатая теория фонологии

* ИЯ АН СССР, Москва.

имеет двойное значение: во-первых, она может служить теоретической базой для особой области фонологических исследований, которую я называю экспериментальной фонологией; во-вторых, она может служить теоретической основой для построения блоков фонологического кодирования в порождающих грамматиках.

DISCUSSION

Martinet:

I do not think there is any need to give the syllable phonological status unless it receives it from accent. There is no need to reckon with syllables in French. I don't even know whether I pronounce *buée* with one or two syllables. The combination of phonemes can be dealt with without recourse to the notion of syllable.

Vanek:

1. How does your phoneme and phonemoid differ from Chomsky's and Halle's systematic phoneme and phonological derived form?
2. If you are operating with three levels, what are the features of the morphophonemic component and how do they differ from the phonological distinctive features?
3. How can you substantiate your claim that morphonemic features differ from the phonological distinctive features and why should they differ?

VERGLEICHENDE UNTERSUCHUNGEN ÜBER DIE SATZINTONATION IN DEUTSCHEN DIALEKTEN

HANS-JOACHIM SCHÄDLICH—HEINRICH ERAS*

Die Untersuchung der Intonationsverhältnisse in den deutschen Dialekten blieb bisher auf einzelne Orte oder kleinere Gebiete beschränkt. Da diese Arbeiten methodisch und materialmäßig uneinheitlich und somit nicht direkt vergleichbar sind, gestatten sie trotz einer Fülle wertvoller Einzelerkenntnisse keine systematische Übersicht. Der vorliegende Vortrag stellt den Versuch dar, zu einer methodisch einheitlichen Beschreibung der Intonationsformen eines bestimmten Satzstrukturtyps für das gesamte Gebiet der DDR beizutragen. Damit soll die großräumige Betrachtung der Intonationsverhältnisse im Deutschen im Sinne eines „Intonationsatlas“ gefördert werden.

Als Grundlage der Untersuchung dienten systematisch durchgeführte Tonbandaufnahmen im Gesamtgebiet der DDR (440 Aufnahmeorte).¹ Voraussetzung unserer Untersuchung war Vergleichbarkeit des Materials.² Sie wurde angestrebt durch Verwendung einer hochdeutschen Textvorlage, die von den Gewährspersonen in den jeweiligen Dialekt übersetzt wurde; durch Einhaltung annähernd gleicher Aufnahmebedingungen; durch Auswahl von 1—3 im Aufnahmeort geborenen dialektfesten Sprechern, die zum Zeitpunkt der Aufnahme älter als 55 Jahre waren. Die überwiegende Zahl der Gewährspersonen entstammte bäuerlichen und handwerklichen Kreisen. Somit kann in Bezug auf diejenigen Merkmale, „die berichten, wie etwas gesagt wird“, als auch in Bezug auf „soziolektale“ Merkmale³ ausreichende Vergleichbarkeit angenommen werden (cf. auch⁴ und⁵).

In dem Satz „Willst du nicht mit mir gehen?“ (Fragesatz mit Spitzenstellung des Verbs und Negationspartikel, sog. Entscheidungsfrage) wurden der Tonhöhenverlauf

* Deutsche Akademie der Wissenschaften zu Berlin.

¹ H.—J. Schädlich und H. Eras: Deutsche Dialektologie und moderne Tonaufnahmetechnik. *Spektrum, Mitteilungsblatt für die Mitarbeiter der Deutschen Akademie der Wissenschaften zu Berlin* 10: 375—382 (1964).

² O. von Essen: Melodische Bewegung der Sprechstimme in deutschen Mundarten. *Forschungen und Fortschritte* 16: 63—64 (1940).

³ G. Hammarström: Können Lautveränderungen erklärt werden? *Proceedings of the Fifth International Congress of Phonetic Sciences*; 336—339 (Basel/New York 1965).

⁴ M. Romportl: *Zvuková stránka souvislé řeči v nářečích na Těšínsku* (Ostrava 1958).

⁵ F. Daneš: *Intonace a věta ve spisovné češtině* (Praha 1957).

und die Position der Hauptakzentsilbe auditiv ermittelt. Jede Veränderung der Tonhöhe, die gleich einem Halbton oder größer war, wurde durch einen Tonbruch gekennzeichnet⁶ (S. 27). Die Zahl der Stufen hängt von der Zahl der Tonhöhenveränderungen in jeder Richtung ab. Die Transkriptionen wurden stichprobenweise mit dem Tonhöhenschreiber von H. Eras⁷ überprüft.

Für den untersuchten Satz wurden zwei Intonationsformen des Satzschlusses ermittelt: steigender und fallender Tonverlauf am Satzschluß. Die Lage der Hauptakzentsilbe wurde durch den links vorhergehenden Tonbruch bestimmt: Tieflage der Hauptakzentsilbe folgt aus fallendem, Hochlage aus steigendem Tonbruch. Durch den Tonverlauf und die Lage der Hauptakzentsilbe sind somit folgende vier Typen konstituiert: 1. Steigender Satzschluß/Tieflage der Hauptakzentsilbe (▲); 2. Steigender Satzschluß/Hochlage der Hauptakzentsilbe (△); 3. Fallender Satzschluß/Tieflage der Hauptakzentsilbe (▼); 4. Fallender Satzschluß/Hochlage der Hauptakzentsilbe (▽) (Abb. 1).

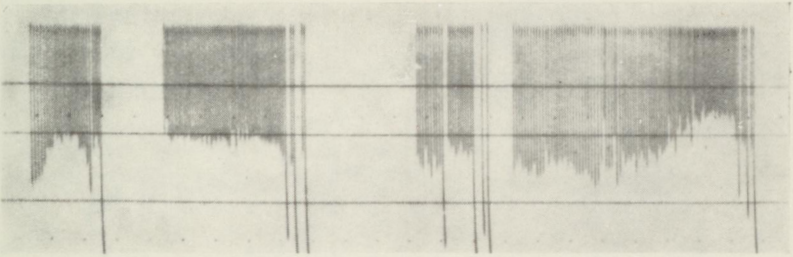
Die genannten vier Typen sind in bestimmter Weise über das gesamte Gebiet der DDR verteilt (Abb. 2). Im Ober- und Mitteldeutschen wie auch im Niederdeutschen tritt jeweils sowohl steigender als auch fallender Satzschluß auf. Ebenso tritt in allen drei Gebieten sowohl Tieflage als auch Hochlage der Hauptakzentsilbe auf. Da der oberdeutsche Gebietsanteil in der DDR gering ist, beschränken wir uns im folgenden auf die Beschreibung der Verhältnisse im Mittel- und Niederdeutschen. Im niederdeutschen Gebiet zeigen nur wenige Prozent der Belege fallenden Satzschluß; im Mitteldeutschen zeigt beinahe die Hälfte der Belege fallenden Satzschluß. Konzentration des fallenden Satzschlusses im Mitteldeutschen tritt in Thüringen auf, während im Erzgebirgischen ausschließlich steigender Satzschluß zu finden ist. Die wenigen Belege für fallenden Satzschluß im Niederdeutschen konzentrieren sich im mittleren und westlichen Mecklenburg. Ohne den Satzschluß zu berücksichtigen, läßt sich über die Lage der Hauptakzentsilbe folgendes sagen: Im Niederdeutschen überwiegen die Belege für Tieflage der Hauptakzentsilbe, im Mitteldeutschen dagegen für Hochlage (in beiden Fällen etwa im Verhältnis 2 : 1).

Zur Vermeidung von Mißverständnissen sei vermerkt, daß die von uns gewählten Parameter nur Grobstrukturen zu erfassen gestatten. Die Feinstruktur z. B. des Typs 1 (Steigender Satzschluß/Tieflage der Hauptakzentsilbe) kann in verschiedenen Gebieten beträchtlich variieren. Durch Berücksichtigung anderer Parameter (Intervallgrößen, weitere Tonbrüche) wäre eine größere Differenzierung möglich.

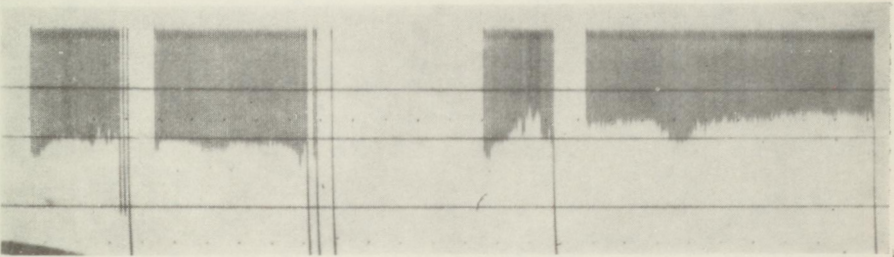
Unsere Beschreibung gestattet Bemerkungen zu der von E. Sievers aufgestellten und später mehrfach erörterten These von den „zwei konträren Generalsystemen

⁶ A. V. Isačenko und H.—J. Schädlich: Untersuchungen über die deutsche Satzintonation. *Studia Grammatica* 7: 7—67 (1966).

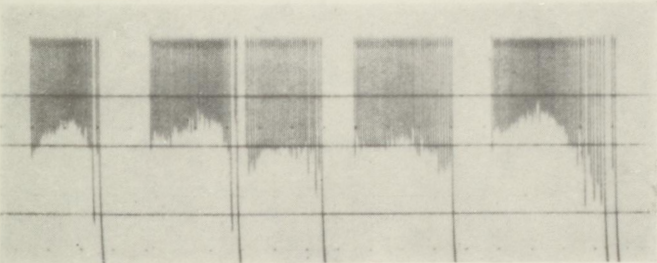
⁷ H. Eras: Die Entwicklung eines Tonhöhenschreibers für phonetische Forschungen. *Proceedings of the Sixth International Congress of Phonetic Sciences*, Prague 1967.



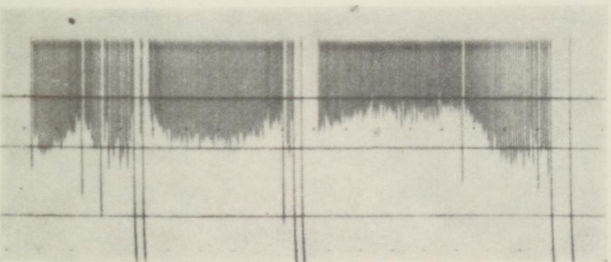
v i l e d u : n i c ' m i d m i : g ä : n



v i l s d u : n i c ' m i d m i : g o : ä n



v i l s d e n i c m e d ' m i : x g i : '



v i l s d a n e m i d ' m i : x j e : n

Abb. 1.

DEUTSCHE
DEMOKRATISCHE REPUBLIK

0 10 20 30 40 km

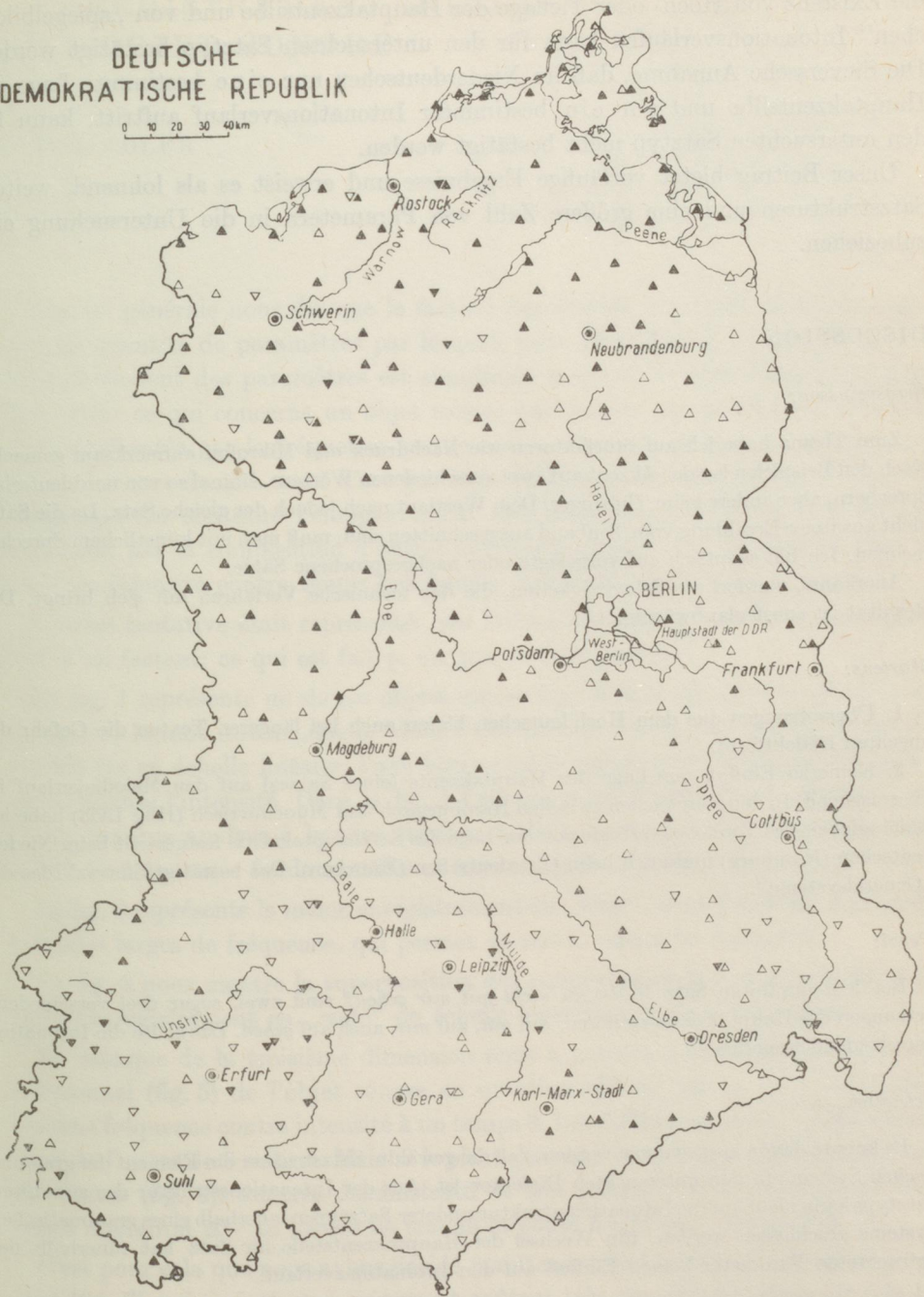


Abb. 2.

der Melodisierung“ im Deutschen.⁸ Die allgemeine Feststellung von Sievers über die Existenz von Hoch- oder Tieflage der Hauptakzentsilbe und von „spiegelbildlichen“ Intonationsverläufen kann für den untersuchten Satztyp bestätigt werden. Die Sieverssche Annahme, daß im Niederdeutschen nur eine bestimmte Lage der Hauptakzentsilbe und nur ein bestimmter Intonationsverlauf auftritt, kann für den untersuchten Satztyp nicht bestätigt werden.

Unser Beitrag bietet vorläufige Ergebnisse und erweist es als lohnend, weitere Satzstrukturen und eine größere Zahl von Parametern in die Untersuchung einzubeziehen.

DISCUSSION

Bergsveinsson:

Zum Thema habe ich auf Störfaktoren wie Nachdruck und Mikrofon aufmerksam gemacht. Nach den Beispielen lag der Akzent auf zwei verschiedenen Wörtern, einmal so von norddeutschen Sprechern, aber anders beim Thüringer. Dem Wortlaut nach jedoch der gleiche Satz. Da die Sätze nicht aus freier Erzählung vom Tonband ausgeschnitten sind, muß man mit künstlichem Sprechen rechnen. Ich bin skeptisch auf vorgelegte oder nachgesprochene Sätze.

Anerkannt wurden die Schwierigkeiten, die das technische Verfahren mit sich bringt. Das Resultat ist sonst glaubwürdig.

Martens:

1. Übersetzungen aus dem Hochdeutschen bieten auch bei längeren Texten die Gefahr der unechten Dialektform.

2. Keinerlei Einfluß der Lage des Hauptakzents (eines Satzes) auf den Melodieverlauf ist überraschend. In den Vergleichen zwischen Hamburgisch und Münchenerisch (Hbg 1953) habe ich Abhängigkeit des Tonhöhenverlaufs von der Lage des Satzhauptakzents festgestellt beim Niederdeutschen (Hamburg) und auch beim Oberdeutschen (München). Das bestätigt Sievers' Idee der „Generalsysteme“.

Penzl:

Bei dem gewählten Satz *Willst du nicht mit mir gehen?* sind zwei, sogar drei verschiedene Satzungen des Hauptakzents zu hören, auf *mit*, auf *mir*, auch auf *gehen*. Das sollte die Intonation entscheidend beeinflussen.

Schädlich—Eras:

Es konnte davon ausgegangen werden, daß die gewählte Satzstruktur ein Element der grammatischen Systeme in den untersuchten Dialekten ist. Aus der Intonationsstruktur des gewählten Satztyps kann nicht auf die Intonationsstruktur anderer Satztypen innerhalb eines grammatischen Systems geschlossen werden. Ein Wechsel der Hauptakzentstelle im Satz hat innerhalb der betrachteten Parameter keinen Einfluß auf den Intonationsverlauf.

⁸ E. Sievers: *Rhythmisch-melodische Studien* (Heidelberg 1912).

L'ENREGISTREMENT TRIDIMENSIONNEL DE L'OBJET SONORE

O. SCHINDLER

Une loi générale nous dit que le facteur connaissance est directement en rapport avec la quantité de paramètres par lesquels nous identifions un objet. De même si l'enregistrement des paramètres est simultané, notre tâche sera d'autant plus facilitée. Pour ce qui concerne un objet sonore quelconque, les paramètres principaux sont représentés par la fréquence, par le temps et par l'intensité. L'enregistrement de ces facteurs deux à deux n'est pas trop difficile; il est bien connu et aussi bien développé. Au contraire l'enregistrement simultané des trois facteurs a été tenté seulement par la composition d'autres enregistrements: p. ex. la superposition des analyses intensité contre temps par bandes d'octaves ou de tiers d'octave.

L'autre tentative était représentée par l'enregistrement qualitatif ou semi-quantitatif d'un facteur: ce qui est fait p. ex. par les sonogrammes.

La fig. 1 représente un de ces objets sonore (qui sera le même pour toute la série de figures) le mot italien „ainole“ dont l'axe vertical représente les fréquences de 0 à 8 000 Hz en échelle linéaire, l'axe horizontal le temps et le noircissement plus ou moins intense l'intensité. Dans la fig. 1 l'enregistrement est fait en „narrow“, c'est à dire par une analyse à bandes étroites de fréquence, qui nous permet de suivre mieux les fréquences fondamentales et harmoniques et la ligne de mélodie.

La fig. 2 représente le même enregistrement en „wide“, c'est à dire par une analyse à bandes larges de fréquence, qui permet de suivre mieux les formants.

La fig. 3 nous montre la superposition des enregistrements 1 et 2 et la fig. 4 présente l'enregistrement en „wide“ en échelle logarithmique des fréquences.

La manque de la troisième dimension nous a poussés à faire un plastique tridimensionnel (fig. 5) de l'objet sonore en question obtenu par la mise ensemble de sections fréquence contre intensité à un temps 0. Les différentes sections sont séparées l'une de l'autre de 8 milisecc. Mais si le plastique, que nous avons photographié aussi par dessus (fig. 6), est très démonstratif et très précis, il comporte une partie de temps d'environ 20—25 heures de travail et coûte à peu près 100 dollars.

C'est pour cela que nous avons décidé maintenant de faire des tracés stéréographiques (fig. 7) qui se font en 1 minute et coûtent très peu et nous donnent très bien la troisième dimension par le système de courbes isosonores qui sont séparées l'une de l'autre de 5 décibels. Ces courbes (dont à la fig. 8 nous avons dissocié les contours

de noircissements et à la fig. 9 nous présentons la réalisation en échelle logarithmique) nous permettent des perspectives très intéressantes soit dans le domaine de la physiologie que dans le domaine de la pathologie que de la synthèse.

DISCUSSION

Grobber:

1. De quelle manière faites-vous ces enregistrements; de toute façon cela ne peut pas être facile.
2. Est-il intéressant d'en faire enregistrer le son?

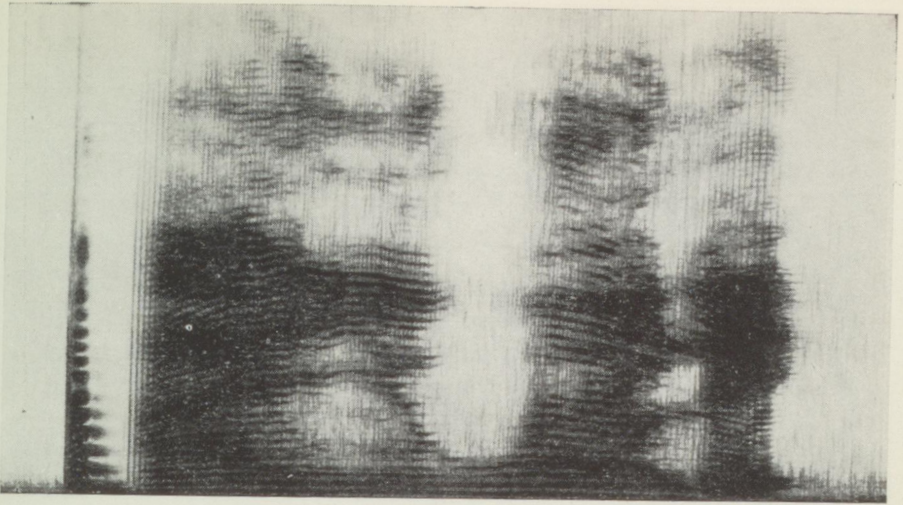


Fig. 1.

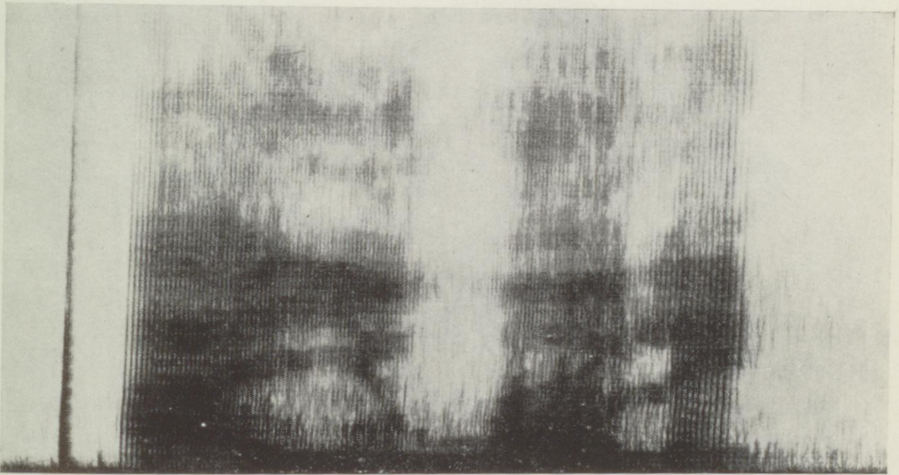


Fig. 2.

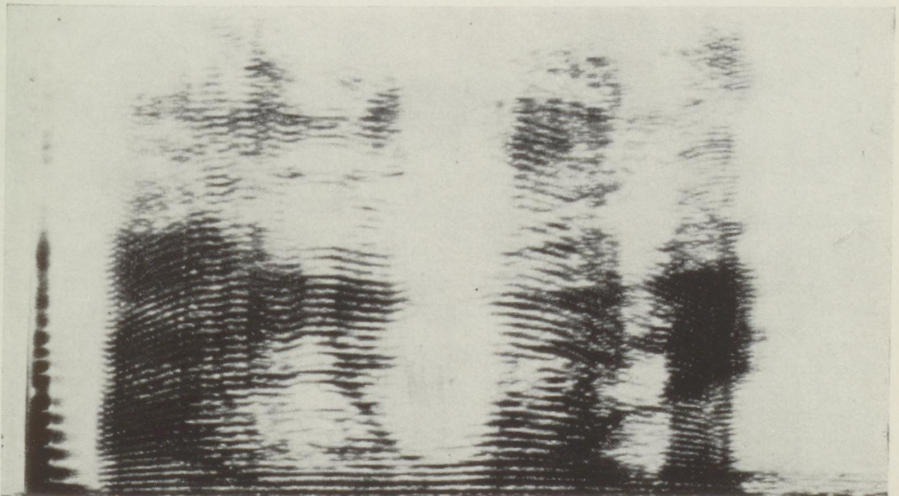


Fig. 3.

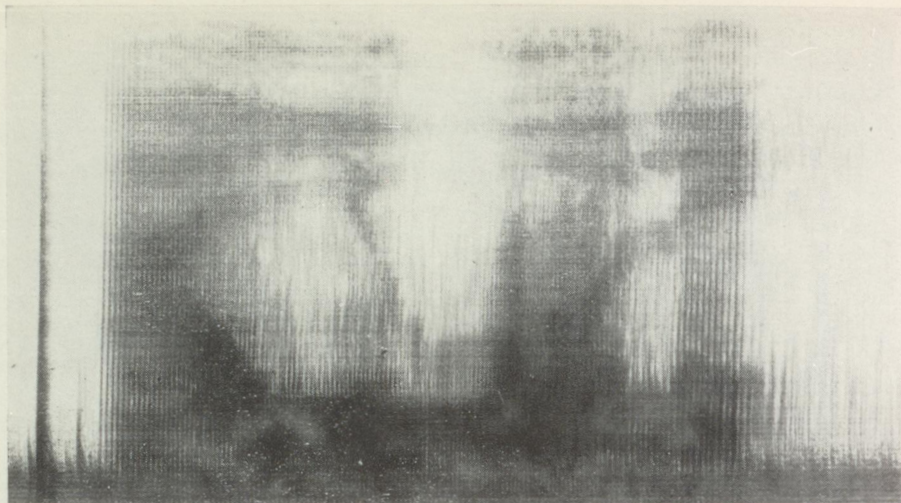


Fig. 4.

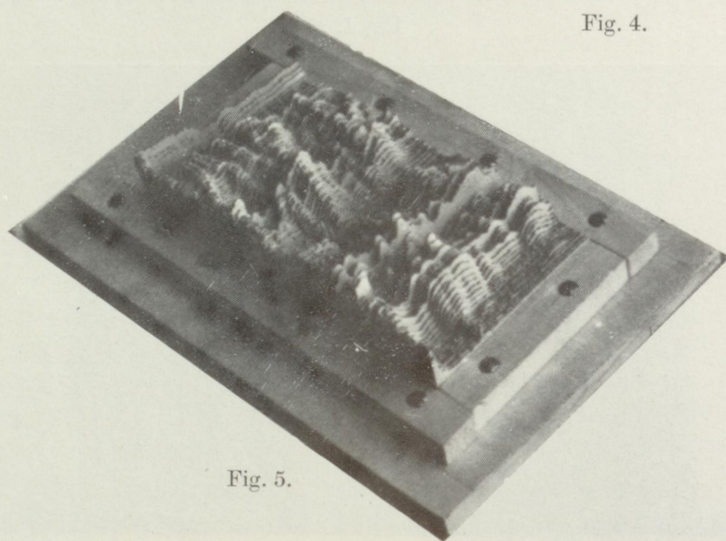


Fig. 5.

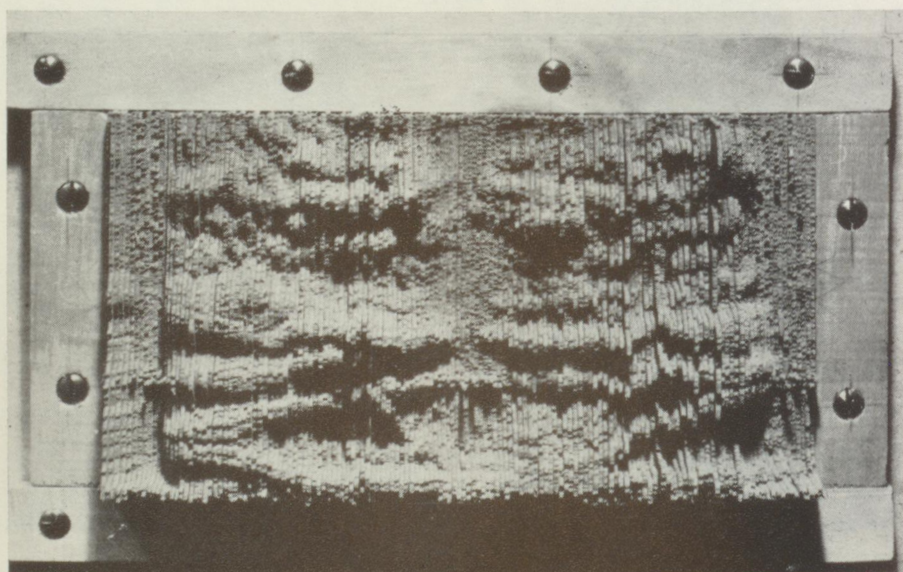


Fig. 6.

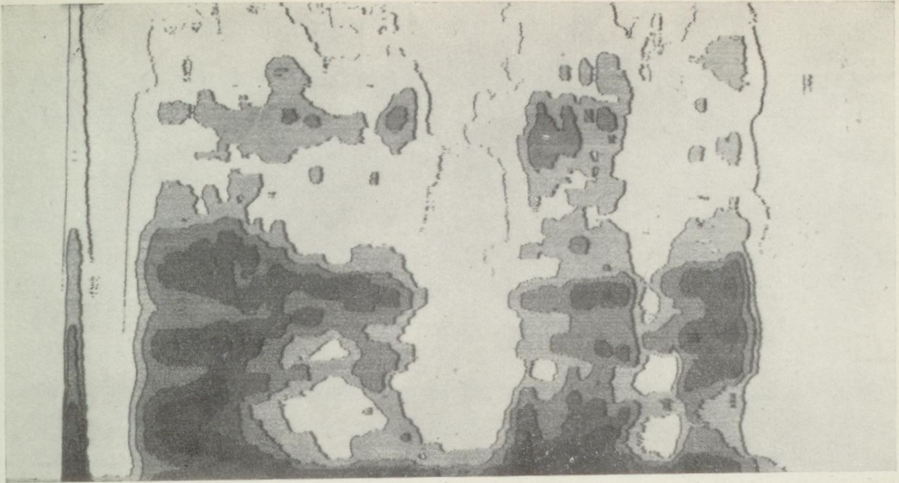


Fig. 7.

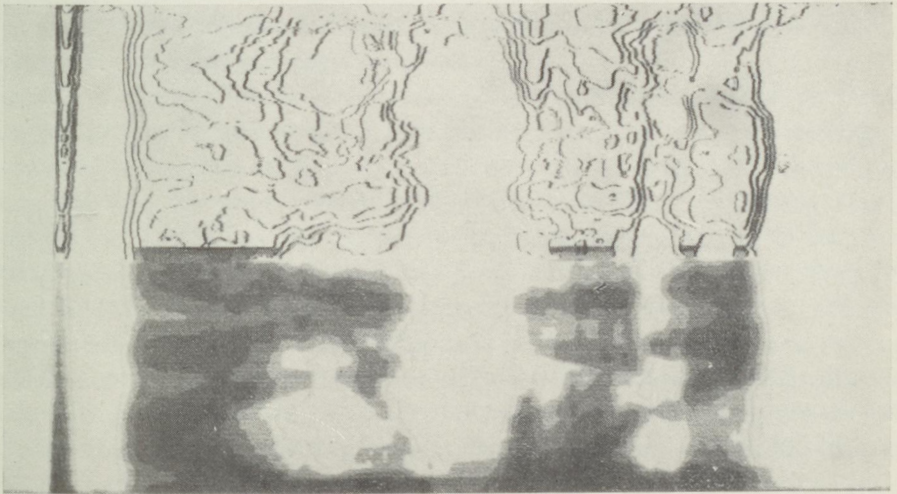


Fig. 8.



Fig. 9.

SUVAG-LINGUA UND DIE PHONETISCHE KORREKTION

ALFRED SCHNEIDER

In der modernen phonetischen Forschung hat sich erst seit einigen Jahren allmählich die Überzeugung durchgesetzt, daß die für die Wahrnehmung der Sprachlaute charakteristischen Komponenten nicht in ihren Klangspektren zu suchen sind, sondern auf der Ebene der Perzeption. Es ist nicht zufällig, daß B. Malmberg an einer neulich abgehaltenen Tagung französischer Phonetiker und Akustiker erklärt hat, in der Phonetik sei das Interesse im Begriff sich zu verlagern zum Verhältnis zwischen dem Stimulus und der Perzeption.

Niemand wird natürlich die Resultate der akustischen Phonetik bestreiten wollen, die in den letzten Jahrzehnten die phonetische Forschung auf die Grundlage der akustischen Bestimmung der Sprache gestellt hat, aber es kann auch nicht gelehrt werden, daß uns die physikalische Analyse der Sprachlaute nur wenig über die Perzeption der Laute und der Sprache schlechthin besagt: ein dt. Wort, dessen Spektrogramm eine ideale Verteilung der Frequenzen und ihrer Intensitäten (Formanten) zeigt, wird auch unter den besten Transmissionsverhältnissen von einem Ausländer, der Deutsch nicht kann, fehlerhaft gehört und wiedergegeben. Auch mit bloßem Ohr wird man mühelos feststellen können, daß der Ausländer nicht nur vieles einfach nicht gehört und folglich ausgelassen hat, sondern daß er in gewissen Fällen bei der Wiedergabe des Gehörten manches hinzugefügt hat, was im Modell physikalisch gar nicht vorhanden gewesen ist.

Im Centre Universitaire in Mons haben wir im Fremdsprachenunterricht innerhalb von 4 Jahren insgesamt über 500 Personen der phonetischen Korrektur nach der verbo-tonalen Methode von Professor P. Guberina unterzogen. Zu diesem Zweck haben wir auch die nötigen verbo-tonalen Apparate verwendet, namentlich das Gerät SUVAG-Lingua. Als Versuchspersonen dienten vor allem Studenten unserer Dolmetscherschule und Schüler zahlreicher Fremdsprachenkurse. Einem überaus aufschlußreichen Experiment mit einem 13-jährigen, schwerhörigen Mädchen, dem im Laufe dieses Jahres Unterricht in deutscher Sprache erteilt wurde, messen wir in diesem Rahmen eine besondere Bedeutung zu. Es ist zu betonen, daß es sich hier um einen schweren Fall von Schwerhörigkeit handelte. Zum Zeitpunkt der Aufnahme des Deutschunterrichtes war sein Gehörvermögen durch Rehabilitierung nach der verbo-tonalen Methode in weitem Ausmaß entwickelt und die Artikulation der frz. Muttersprache zum großen Teil eingeübt, die Aussprache der einzelnen Laute wies aber

noch zahlreiche Mängel auf. Die deutsche Sprache sollte nun durch ihr neues, der Schülerin vollkommen unbekanntes Lautsystem, neue Intonationsarten und andersartigen Rhythmus das Gehör des Mädchens in völlig neue Hörverhältnisse setzen, seinem Gehirn neue und reichere Möglichkeiten der Auswahl von Frequenzen und Intensitäten gewährleisten, um ihm akustisch-psycho-physiologisch die Perzeptionsmöglichkeiten der Muttersprache zu erweitern.

Die sowohl im Fremdsprachenunterricht wie auch im Deutschunterricht mit dem schwerhörigen Mädchen gesammelten Erfahrungen in der phonetischen Korrektur haben die Richtigkeit der Grundsätze der verbo-tonalen Methode bestätigen können. Die Ergebnisse haben gezeigt, daß die Sprache auf der Grundlage begrenzter Frequenzbänder gehört wird, in denen der Diskontinuität der Frequenzen, Intensitäten und des Faktors Zeit eine entscheidende Bedeutung zukommt. Unsere schwerhörige Schülerin hat die Laute der dt. Sprache in der Diskontinuität zweier für sie optimaler Frequenzbänder besser gehört und verstanden als im gesamten Frequenzbereich; die Studenten und Schüler der Fremdsprachenkurse perzipierten die Laute der verschiedenen Fremdsprachen in relativ schmalen aber optimalen Frequenzbändern gut, im gesamten Frequenzbereich haben sie aber die Laute durch nächstliegende Laute der Muttersprache vertauscht. Daraus kann man schließen, daß das Gehör tatsächlich funktioniert aufgrund einer Auswahl von einigen strukturalen Elementen und nicht linear und quantitativ.

Es besteht aber nicht nur eine gegenseitige Abhängigkeit zwischen dem Hören und dem Sprechen, was besagen soll, daß es äußerst schwierig ist, etwas auszusprechen, was man nicht hört bzw. nicht gut hört, sondern es besteht auch gleichzeitig eine gegenseitige Abhängigkeit zwischen dem Sprechen und dem Hören, was wiederum bedeutet, daß man etwas nur schwer hören kann, was man nicht selbst aussprechen bzw. nachsprechen kann. Der Schüler hört die Fremdsprache aufgrund des Systems seiner Muttersprache, und es hat sich gezeigt, daß z. B. Frankophonen ein dt. [h], wie in „Haus“ in der Regel überhaupt nicht hören, daß sie ein engl. [θ], wie in „think“, wohl hören aber als ein [s] aussprechen. Dabei hat sich ergeben, daß diese Fehler ein jeder Frankophone macht, d. h. daß die Fehler einem System folgen. Ein gutes Gehör allein genügt also nicht, um einen Sprachlaut richtig zu hören, sondern man muß auch gut artikulieren können, damit man richtig hören kann. Um das zu erreichen, muß man den Lernenden verbessern, so lange verbessern, bis seine Aussprache fehlerlos geworden ist.

Nach unserer Erfahrung glauben wir behaupten zu können, daß die verbo-tonale Methode zu gewissen objektiven Ergebnissen geführt hat, die besonders im Fremdsprachenunterricht und in der Rehabilitation der Schwerhörigkeit von großer Bedeutung sind. Das gilt vor allem für die Erkenntnis, daß Sprachlaute von der Struktur aus untersucht werden sollen, daß der Auswertung der wirksamsten aber quantitativ begrenzten Elemente der Laute eine besondere Bedeutung zukommt, daß das Sprechen und das Hören strukturalen Gesetzen unterliegen, und daß es eine gegenseitige Abhängigkeit gibt zwischen Sprache und Perzeption der Sprache.

LITERATUR

1. Guberina, P., *The Verbotonal Method and its Application to the Rehabilitation of the Deaf, Proceedings of the International Congress on Education of the Deaf, Washington 1964.*
2. Guberina P.—Gospodnetić J., *Audition et articulation à la lumière de la méthode verbo-tonale, Proceedings of the 12th International Speech and Voice Therapy Conference, Padua 1962.*

CHARAKTERISTISCHE SCHALLNACHAHMENDE LAUTGRUPPEN IM DEUTSCHEN

VILIAM SCHWANZER

Wenn man sich P. Trosts¹ scharf formulierte und sichtlich auch scharf gemeinte Bemerkungen über alles das, was mit Phonosemantik, Lautbedeutungslehre oder ähnlichem zusammenhängt, ins Gedächtnis ruft, erscheint es bedenklich, neuerdings die Frage nach Zusammenhängen zwischen Lautung und Bedeutung aufzuwerfen.

Es soll auch nicht auf die in der Tat „mystisch-seherischen“ Weissagungen M. Rubinyis² von der Zukunft der Lautnachahmungstheorie für die Sprachwissenschaft, die stellenweise überschwenglichen und ziemlich einseitigen Betrachtungen W. Wißmanns³ oder die mehr defensiv eingestellte zusammenfassende Darstellung E. Brocks⁴ zurückgegangen werden. Wir wollen vielmehr das Problem von einer anderen Seite angehen und die Grundfrage so stellen:

Gibt es unterhalb der Morphemgrenze gewisse rekurrente, d. h. in bestimmten Morphemen oder Wörtern immer wieder vorkommende Lautkonstellationen oder — kombinationen, aufgrund derer eine linguistisch relevante, auf die Semantik mitbezogene Klassenbildung möglich wäre?

Wir gehen also nicht a priori von einer „Bedeutung“ der Laute aus. Außerdem respektieren wir den Grundsatz, daß die Laute (Phoneme) zwar eine bedeutungsbildende und bedeutungsdifferenzierende Funktion, aber keine Eigenbedeutung haben und erst in der Kombination — in einigen Fällen wohl auch mit einem phone-misch-phonetischen Null-Element — fungieren können.

In der traditionellen Literatur verfuhr man in der Regel anders: Aus einigen wenigen Wörtern oder Wortteilen, von denen man annahm oder behaupten zu können glaubte, daß sie schallnachahmend seien, hob man einzelne Laute oder bestimmte Lautsequenzen (clusters) heraus und sah in diesen den bedeutungstragenden Kern des Morphems (Semems), der als Nachahmung natürlichen Schalles

¹ O „významu“ hlásek. (Über die „Bedeutung“ der Laute.) *Studie a práce lingvistické I*, Praha 1954, 67—72.

² Das Problem der Lautnachahmung. *GRM* 5 (1913), 497—512.

³ Ausdrucksworte und Lautverschiebung. *ZfdAuDL*, 76. Bd., 1939, 1—12.

⁴ Der heutige Stand der Lautbedeutungslehre. *Trivium* 2, 1944, 199—219.

interpretiert und dem distinktive Kraft zugeschrieben wurde.⁵ Man schenkte hiebei den auffallenden, ungewöhnlichen Sequenzen besondere Aufmerksamkeit, weil man meinte, daß zwischen konventioneller Lautung und Schallnachahmung ein wesentlicher Unterschied sein müsse. Es wurde nicht bedacht, daß schließlich auch die Nachahmung natürlichen Schalles durch artikulierte Laute niemals — auch nicht durch außergewöhnliche Bündelungen der nun einmal in jeder Sprache vorhandenen Laute (Phoneme) — vollkommen sein kann,⁶ daß vielmehr die Grenze zwischen konventioneller Lautung und Annäherung an den Naturlaut (in wirklichen Schallwörtern wie *gluck*, dazu *glucken*, *gluckern*, *glucksen* usw., *puff*, dazu *puffen*, *Puffer* usw., *klitsch*, *klatsch* u. a.) sehr verschwommen ist und letzten Endes auch diese Annäherung eine Konvention darstellen kann: Man entnimmt der Lautung, welcher natürliche Schall angedeutet wird, aber nur dann, wenn man den Usus und die Übereinkunft kennt, wonach in der betreffenden Sprache der oder jene natürliche Schall wiedergegeben, bzw. bezeichnet wird. Also auch diese Wörter bezeichnen einen Schall und sind keine direkte Reproduktion des Schalles, auch in ihnen ist die Form nicht durch den Inhalt (die Realität) motiviert. Nicht der Schall wird in die Rede eingebaut, sondern ein Wort als konventionelles Zeichen für diesen Schall. Deshalb sind Nachahmungen des gleichen Schalles in verschiedenen Sprachen meist verschieden, z. B. d. *pitsch*, *patsch*, *plan[t]sch*; slow. *člup[s]*, *žblnk*; tschech. *plác*, *plešk* (Fall oder Sprung ins Wasser). Auch in expressiven Wörtern gilt dies: d. *au* (Schmerzruf); slow. *jaj*, *jáj*, *jajaj*; engl. *oh*; franz. *ah*, *ahi*, *ouf*. Wären diese Wörter wirkliche Schallnachahmungen, bzw. allgemein-menschliche Expressiväußerungen, so müßten sie eigentlich in allen Sprachen gleich, also international sein.

Als Beispiel, wie bei der eingangs erwähnten Fragestellung vorzugehen wäre, seien die deutschen Wörter mit den Initialsequenzen *gn*- und *kn*- (zwei anzahlmäßig beschränkte Lautungsklassen) angeführt. Sie können jede für sich in Unterklassen geteilt werden, wobei in jeder Unterklasse nur semantisch zusammengehörige Wörter vorkommen. Es zeigt sich, daß von 44 Wörtern mit den Anfangslauten *gn*- 14 Wörter (31,8%), z. B. *gnägeln*, *gnatzen*, *gniäten*, *gnietsch*, *gnöckern*, den Bedeutungskern „Unlust, Unzufriedenheit“, zwei davon mit der Modulierung „Geiz, Neid“ (*Gniefke*, *gnittschäwsch*) enthalten, weitere 2 Wörter (4,5%), nämlich *gnappen*, *gnagen*, „Mühe, Anstrengung“ in ihrem Bedeutungskern bergen, 5 Wörter (11,4%), u. zw. *gnübbeln*, *gniegeln*, *Gniegelstein*, *gnieseln*, *gnieweln*, auf „Reiben, Drücken“ hinweisen, 2 Verben (4,5% der Wörter) „gezielte Destruktion, Tätlichkeit“ ausdrücken (*gnepfen*, *gnübbeln*) und 16 Wörter (36,4%) — darunter manche mit nicht ursprünglicher Sequenz *gn*- (*Gnade*, *gnädig*, *Gneiß*) oder mit aus fremden Sprachen

übernommener Lautung (*Gnom*, *Gnu* aus *ngu*) verschiedener Bedeutung sind. Schallbezeichnend sind 5 Wörter (11,4%): *gnaschen*, *gnassen*, *gniedeln*, *Gniedelstein*, *gnütern*.

Von 132 *kn*-Wörtern bedeuten „Unlust, Unzufriedenheit, Unzulänglichkeit“ 11 Wörter (8,3%), z. B. *knauern*, *knaufen*, *knaunzen*, *knapp*, *Knauser*, 11 Wörter (8,3%) enthalten „Mühe, klein[lich]e Tätigkeit“, z. B. *knabbern*, *knappern*, *knarfeln*, *kneifeln*, 11 Wörter (8,3%) „Reiben, Drücken“, z. B. *Knatsch*, *kneifen*, *kneten*, *knutschen*, 20 Wörter (15,1%) weisen auf „Härte, Verhärtung, Festigkeit“ hin, z. B. *Knagge*, *Knast*, *Knüttel*, *Knochen*, 14 Wörter (10,7%) auf „Zusammenballung“, etwa *Knäuel*, *knäueln*, *Knödel*, *Knolle*, *Knoten*, *knüpfen*, auf „Zorn, Heftigkeit“ 6 Wörter (4,6%), nämlich *knarsch*, *Knergel*, *knergeln*, *Knies*, *knirb[e]lig*, *knotteln*, und auf „destruktive Tätigkeit, Destruktion“ 7 Wörter (5,3%): *knäseln*, *knick[s]*, *knick[s]en*, *Knick[s]er*, *knittern*, *Knitter*, *knitt[e]rig*. Verschiedene Bedeutung haben 35 Wörter (26,5%), manche übernommen oder mit sekundärer Lautfolge *kn*- Schallbezeichnend sind 17 Wörter (12,9%), z. B. *knack[s]*, *knack[s]en*, *Knall*, *knallen*, *knattern*, *knispeln*, *knören*.

Es ergeben sich Klassen von Wörtern mit gewissen Gemeinsamkeiten der Lautung, die in bestimmte semantische Bereiche gehören, keineswegs jedoch als Schallimitationen oder von der Bedeutung her als lautlich motiviert angesehen werden können. Soweit ist m. E. derzeit der Zusammenhang zwischen Lautung und Bedeutung wissenschaftlich vertretbar. Es kann vom Lautformalen als Primärkriterium und vom Semantischen als sekundärem Kriterium her eine umfassende sprachbezogene Klassifizierung des Wortschatzes geschaffen werden.

⁵ Vor diesem Irrtum warnt W. Haas, *Two Types of Phonological Value. Proceedings of the Fourth Int. Congress of Phonetic Sciences* (1961), The Hague, 1962, 625—629. Er spricht — in Anlehnung an J. R. Firth — von *phonaesthetic elements*.

⁶ Vgl. W. Mues, *Vom Laut zum Satz*. Heidelberg 1964, 27 („akustische Parallelen“).

VOWEL QUALITY AND WAVE SHAPE

CELIA SCULLY

Many acoustic studies have been made of vowels and the frequencies F_1 and F_2 of the first and second formants are generally used as parameters for vowel quality. There are several objections to these measures however. Firstly, similar vowel sounds made by men, women and children have very different formant patterns. In the case of [æ] for example, a child's first formant can be about 50% higher than a man's.¹ Some method of normalizing formant data seems essential. The ratio of F_2/F_1 seems to be a better measure for front vowels than the absolute values of F_2 and F_1 , but this is not true of back vowels.² Some overlap occurs in the F_1 versus F_2 plot. For example a man's [a] and a child's [ɔ] can have the same values for F_1 and F_2 ; by changing the fundamental frequency but not the formant frequencies the vowel quality can be changed.³ Secondly, it is not always easy to separate adjacent formants or indeed, for some back vowels, to find spectral peaks at all. This is because the resonances of the vocal tract do not always result in clearly separated peaks in the frequency spectrum. Thirdly, an F_1 versus F_2 plot takes no account of the relative strengths of the formants. In close front vowels F_2 is arbitrarily selected from the several strong high formants, of which F_2 is not always the strongest; and it has been shown that the relative amplitudes of the second and third formants can influence the perceived vowel quality.⁴

Although different speakers probably articulate similar vowels in a similar way, thus producing their characteristic formant patterns, it does not necessarily follow that perception of these sounds proceeds in the same way. Indeed, formantless complex sounds can be identified consistently as different vowels.⁵ It may be that

¹ G. Fant: „A Note on Vocal Tract Size Factors and Non-Uniform F-Pattern Scalings“. Royal Inst. of Tech. (Stockholm) *QPSR-4/1966* pp. 22—30.

² G. E. Peterson: "Parameters of Vowel Quality". *J. of Speech and Hearing Res.* 4 (1961) pp. 10—29.

³ R. K. Potter and J. C. Steinberg: "Toward the Specification of Speech". *J. of the Acous. Soc. of Amer.* 22 (1950) pp. 807—820.

⁴ R. W. Peters: "Dimension of Quality for the Vowel [æ]". *J. of Speech and Hearing Res.* 6 (1963) pp. 239—248.

⁵ A. Carpenter and J. Morton: "The Perception of Vowel Colour in Formantless Complex Sounds". *Lang. and Speech* 5 (1962) pp. 205—214.

in speech perception the pitch of the voice producing a vowel sound is taken into account by the listener in assessing vowel quality, but in the experiment reported here the aim was to look for purely acoustic measures of vowel quality; ones not based on knowledge of the speaker's physiology.

Vowel sounds of Southern English were made by a few speakers imitating the experimenter. 62 of the sounds were made into 200 mS segments and presented in random order to 10 phonetically trained listeners, who were asked to identify each stimulus with one or more phonetic symbol, using English Received Pronunciation as a guide. [i], [æ], [a] and [u] were identified with nearly 100% agreement. Formant frequencies were measured for the sounds and, in addition, detailed records of the wave shapes or oscillograms, were made, using an Ultra-Violet Recorder running at 100 ips and responding up to 7 000 Hz without significant distortion of the wave shape.

Fig. 1. shows the wave shapes for four sounds intended as [i] and identified as such by at least 9 out of the 10 listeners. Two prominent oscillations, or ripples, can be distinguished easily and their frequencies measured. The slower ripples have frequencies between 260 Hz and 350 Hz, corresponding to the first formants. The faster ripples have frequencies ranging between 3 030 Hz for a man and 3 640 Hz for the child. That is, near or a little above the frequency of F_3 for the men, between F_2 and F_3 for the woman and near or a little above F_2 for the child. The spread in ripple frequencies is very much smaller than the spread in F_2 frequencies, as table 1. shows.

Table 1.

Speaker	Ripple Frequencies			Formant Frequencies	
	F_L	F_H	F_H/F_L	F_2	F_2/F_1 (approx.)
[i]					
RS (Child)	352 Hz	3640 Hz	10.3	3600 Hz	10.3
CS (Woman)	272	3270	12.0	2900	11.6
PADM (Man)	266	3180	11.9	2250	7.5
WRBA (Man)	273	3030	11.1	2200	7.3

Fig. 2. shows that, in some cases at least, this high frequency ripple is not merely the strongest of the high frequency formants, but is an amplified and damped wave resulting from two nearby prominent harmonics. Here the waveshapes of three [i] sounds are shown after high-pass filtering at 1 500 Hz. The two children's wave shapes show beating between two adjacent harmonics and the man's wave shape shows beating between harmonics separated by four times the fundamental fre-

quency. The frequency agrees well with results for two-formant synthetic vowel sounds.^{6 7}

Frequency and time measures are two aspects of the same thing; a formant shows as a frequency band in a spectrogram or as a ripple in the wave shape. But the wave shape, unlike the spectrogram, reveals combinational features of the formants.

For [æ] there is one prominent ripple at about 1 900 Hz. This is between F_1 and F_2 for the children, but between F_2 and F_3 for the adults. Back vowels showed ripples decreasing in frequency down to [u]. [ɔ] and [a] were well separated by their ripple frequencies although their F_1 and F_2 frequencies overlapped.

These results tend to support the view that front vowels are perceived as two-formant sounds, back vowels as one-formant sounds with [æ] or [a] at the boundary. Whether the measure of vowel quality should be absolute frequencies or ratios of frequencies, ripples in the wave shape could possibly be more appropriate than formants in the spectrogram as acoustic parameters of vowel quality.

DISCUSSION

Fujisaki:

It is quite true that the Sonagraph is far from an ideal tool for frequency analysis, but I believe one should not denounce the formant concept because of the incomplete instrumentation.

Mainwold:

A short remark to the phase discrimination in the human ear: In masking experiments with several harmonic components with fixed relative phase as masker, it can be shown that the threshold for an additional sinusoidal test-tone (with a frequency lying somewhere in between) is strongly dependent on the phase relation between the spectral components of the masker though the overall loudness of the complex sound does not change. Consequently there seems to be some evidence that the phase information is still present on a high processing level and might, in fact, influence the sound quality.

Merhaut:

What is the influence of phase of formants on the perception (when there is a strong influence on the shape)?

Slis:

Do you find with visual inspection of oscillograms of whispered vowels a differentiation between the different vowels comparable to the results with the normally spoken ones?

⁶ P. Delattre, A. M. Liberman, F. S. Cooper and L. J. Gerstman: "An Experimental Study of the Acoustic Determinants of Vowel Color; Observations on one—and two-formant Vowels Synthesized from Spectrographic Patterns." *Word* 8 (1952) pp. 195—210.

⁷ G. Fant and A. Risberg: "Auditory Matching of Vowels with Two Formant Synthetic Sounds". Royal Inst. of Tech. (Stockholm) *QSPR-4/1963* pp. 7—11.

Ad Merhaut, ad Fant: Phase relationships between the harmonics concerned in high frequency ripples probably are important in determining the amplitude of the resultant ripple. There is now a good deal of evidence to suggest that in auditory perception time measures, perhaps even of fine structure, may be possible. One method of processing the wave form is to make a time measure for half a ripple period at the point in time where the ripple amplitude is a maximum.

Ad Slis: Records for a few whispered vowels have been made. Their oscillograms show characteristic ripples although naturally there is no wave shape periodicity associated with a fundamental as there is for normal voiced vowel sounds. Although as yet no measurements have been made, visual inspection indicates that the ripple frequencies vary for different whispered vowels in the same sort of way that they do for normally spoken ones.

Ad Fujisaki: My objections to the use of formant patterns especially of F_1 and F_2 are not dependent upon the limitations of particular instruments. In any Fourier analysis, phase relationships between the harmonics are not shown.

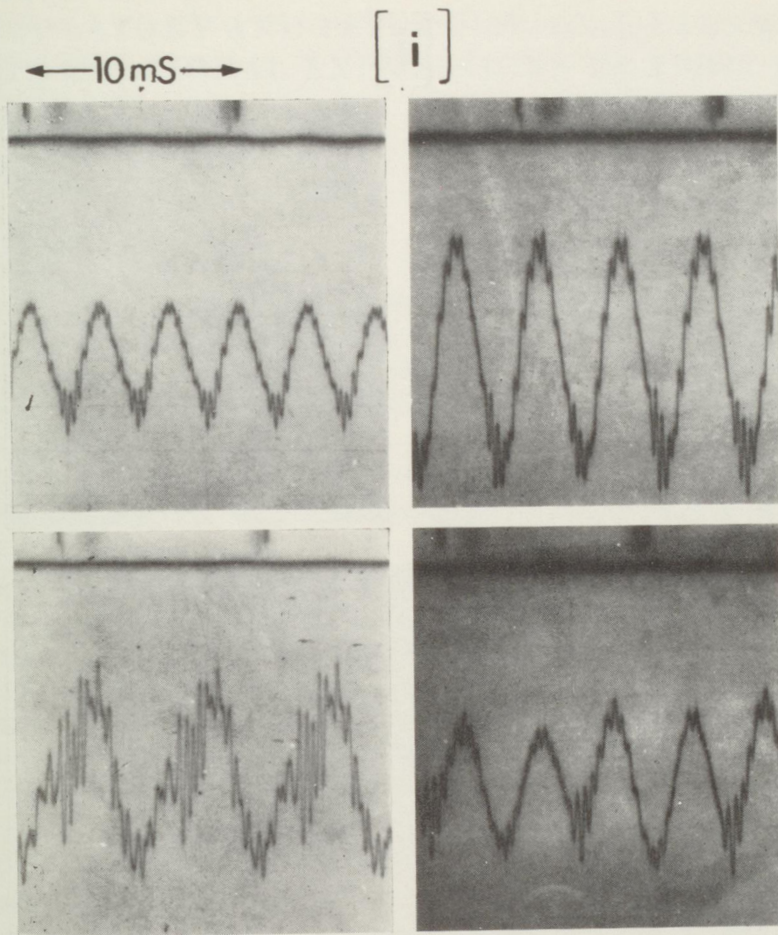


Fig. 1. Speakers: RS (child), CS (woman), PADM (man), WRBA (man).

[i]

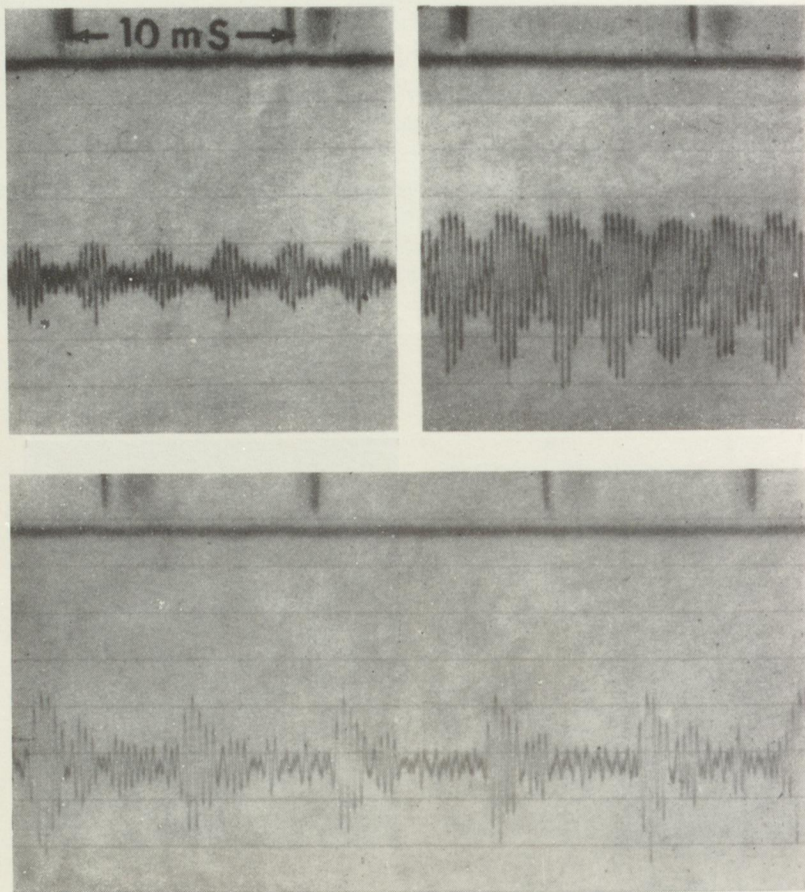


Fig. 2. Speakers: RS (child), Ch. S (child), WRBA (man).

RELATIONS BETWEEN THE ACOUSTIC, ARTICULATORY AND PSYCHOLOGICAL PARAMETERS OF THE EMOTIONAL EXPRESSIONS IN SPEECH

A. SYCHRA — K. SEDLÁČEK

In our study published in *Folia Phoniatrica* in 1963 (15: 89—98) we tried to demonstrate that the expression of joy or grief is closely related to the mean pitch of the phrase. Our supplementary revision of the same material has proved that the correlation index according to the test of Kendall is + 0,66 which is the value significant on the level of 1 %.

Such definite correlations do not occur very often. The constant melodic or melodico-rhythmical patterns are much less frequent. Complicated vectors of acoustic parameters are usually at play.

When studying these vectors we have to be aware of the fact that the speaker communicating emotions is not able to control immediately the acoustic parameters such as the upper limit of the spectrum, that is the highest partial component of the sound, or the frequency and the intensity of the formants. But what he does control are the basic physiological or as the case may be psychological intentions, or let us say instructions, from which in our opinion the most important are: brightness or darkness of the voice, softness or harshness of the articulation, purity or hoarseness of the voice, liveliness or monotony of the melodic or rhythmical movement, smiling or weeping voice etc.

At first the correlations between these instructions and the expression of joy or grief and love or anger were established on the basis of several groups of observers who registered their subjective feelings which were evoked by 23 different recitations of two sentences. Figure 1 demonstrates the positive correlation of the rank-order JOY-GRIEF with the rank-order BRIGHTNESS-DARKNESS of the voice. The correlation index $t = +0,65$ is significant on the 1 % level. A similar positive correlation exists between JOY-GRIEF and PURITY-HOARSENESS of the voice, the coefficient of the correlation being +0,45 (equally significant on the 1 % level).

On the contrary the rank-order LOVE—ANGER is not correlated either with BRIGHTNESS—DARKNESS, or PURITY—HOARSENESS of the voice. There exists however a negative correlation between the rank-order LOVE—ANGER and the rank-order SOFTNESS—HARSHNESS ($t = -0,49$ significant on the 1 % level). No further correlations of this rank-order with the physiological instructions have been found.

Our second task was to try to define the physiological or psychological instructions

in terms of acoustic parameters. We communicate to-day only some of the positive findings of our rather extensive investigations.

First of all we have to emphasize that in accordance with our earlier investigations which we have mentioned, the most important factor is the mean pitch of the voice. To a certain extent, all physiological or psychological instructions we have studied are somehow related to this parameter. A very close correlation exists between the mean pitch and the rank-order BRIGHTNESS—DARKNESS ($t = +0,676$ which is significant on the 1 % level). Among other acoustic factors 3 parameters characterizing the spectrum of the voice are very important, namely: the relative intensity of the higher components of the voice, the upper limit of the spectrum and the factor, we call the upper limit of the periodicity, that is the highest harmonic (periodic) tone.

As far as the intensity of the higher components of the voice spectrum is concerned, the positive correlation of this parameter with the rank-order SOFTNESS—HARSHNESS has been established. For this purpose, the intensity of the 3rd and 4th formant was taken into consideration with the result of $t = +0,328$ for the 3rd formant and $t = +0,359$ for the 4th formant, both significant on the 5 % level.

The material we have chosen for the present communication enables us to differentiate between the instructions of BRIGHTNESS—DARKNESS and SOFTNESS—HARSHNESS. The respective results are demonstrated on the table. The rank-order BRIGHTNESS—DARKNESS shows the positive correlation with the pitch, with the upper limit of the spectrum and the upper limit of the periodicity. No correlation has been found in the intensity of high components. On the contrary, the rank-order SOFTNESS—HARSHNESS is characterized by the negative correlation with the intensity of the higher sound components and the negative correlation of the upper limit of the periodicity, whereas the remaining acoustic parameters are without any correlation. The results obtained are in accordance with our suppositions as to the phonatory mechanisms. We assumed we would find a higher positive correlation between the rank-order PURITY—HOARSENESS and the upper limit of the periodicity which we could not demonstrate on the basis of this material because of the absence of hoarse voices. But we have discovered this dependence in another group of our experiments containing phrases with the expression of FEAR.

These conclusions are naturally valid for the present material only; generalization will be possible only on the basis of more extensive experiments connected with statistical factor analysis. Nevertheless, in our opinion the possibility of defining the emotional expression of speech by means of the acoustic parameter has been thus demonstrated.

DIFFERENZIERUNG DER FORMANTENGEBIETE IN DEN ANFÄNGEN DER STIMMENTWICKLUNG DES KINDES

E. SEDLÁČKOVÁ*

Erst in den letzten Jahren beginnt sich das Studium des akustischen Bildes der Kinderstimme mit Hilfe exakter Methoden der Klanganalyse zu entwickeln. Die bis jetzt veröffentlichten Arbeiten widmeten sich vorwiegend der Feststellung grundlegender akustischer Eigenschaften der Kinderstimme und dies hauptsächlich bei Neugeborenen. Einer systematischen Beobachtung der Entwicklung der Klangfarbe und den Anfängen einer Differentiation der artikulierten Sprache wurde bisher keine genügende Aufmerksamkeit gewidmet.

Aus diesem Grunde richten wir unser Augenmerk in systematischen regelmäßigen Klangaufnahmen auf eine Gruppe von ursprünglich 30 Kindern, die sich später auf 19 Kinder verringert hat und zwar vom Augenblick ihrer Geburt an, bis einsteilen zu ihrem 6. Lebensjahr. In der heutigen Mitteilung wollen wir uns der Entwicklung der Farbe der Stimme und der Selbstlautbildung im Zeitabschnitt von der Geburt des Kindes bis zu den Anfängen der selbstständigen Sprache in den ersten drei Lebensjahren widmen. Wir wollen speziell über die spezifische Farbe der Säuglingsstimme und die dynamische Entwicklung der Lautformanten sprechen.

Auf dem ersten Bild (Abb. 1) ganz oben ist das summarische statistische Spektrum der Stimme von Neugeborenen veranschaulicht. Die Höhe der Säulen stellt die Zahl der Formanten dar, die im entsprechenden Frequenzgebiet vorkommen. Wir sehen hier markant ausgedrückte Maxima der ersten drei Formanten, zwischen denen sich sogar Lücken befinden. Diese hochgradige Uniformität der Schreie von Neugeborenen zeugt dafür, daß es sich bei der Stimmäußerung in diesem Alter bloß um eine einfache reflektorische Bewegung handelt. Die akustische Farbe der Stimme bei Neugeborenen ergibt sich aus den Formanten in der Höhe von 1 400, 2 400 und 3 400 Hz und aus der Abschwächung bis zur Unterdrückung des Grundtones. Wir haben auch Formanten der einzelnen Nuancen der Stimme der Neugeborenen festgestellt, besonders eine dem Laute *a* und dem Laute *ä* ähnliche Färbung, aber die Unterschiede zwischen denselben sind nicht so groß, um das summarische Spektrum grundsätzlich verändern zu können.

Weiter untersuchten wir die Veränderungen der Formanten in den einzelnen

* Phoniatische Klinik der Karls Universität in Prag. Vorstand: Prof. Dr. M. Seeman, DrSc.

Lebensstufen. Die genau abgegrenzten Formanten des Neugeborenenzeitabschnitts erweitern sich im dritten und vierten Monat, sie beginnen zu schwanken, so daß sie auf dem Diagramm kaum angedeutet sind. Dies entsteht dadurch, daß in den Stimmäußerungen der Kinder neue Klänge erscheinen, die auch in jenen Gebieten

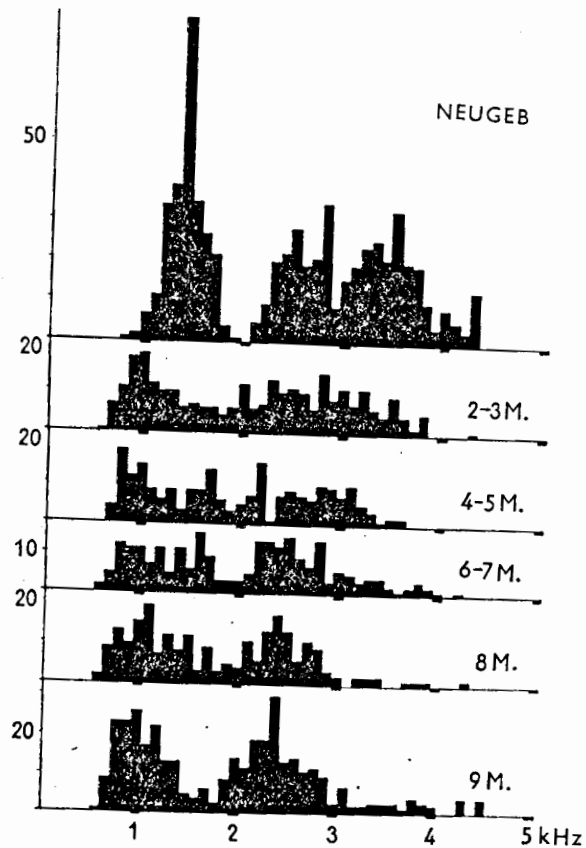


Abb. 1.

Formanten haben, die früher leer waren. Diese Erscheinung schreibe ich dem Einfluß der Beteiligung des Gehörs zu, das sich als akustische Rückkoppelung in die Steuerung der Phonation des Kindes einschaltet. In den weiteren Monaten organisieren sich die schwankenden Formanten wieder in neue Gebiete, die sich den Lautformanten des Sprachmilieus zu nähern beginnen.

In der Periode der Nachahmung und der spontanen Bildung von Wörtern im zweiten Lebensjahr vertiefen sich die Formanten und nähern sich einander einigermaßen.

Im Verlauf des zweiten und des dritten Lebensjahres, wenn die Kinder schon alle fünf Grundvokale der tschechischen Sprache bilden können, konnten wir auch die Bewegung ihrer Formantengebiete bestimmen. Sie sind in den Diagrammen der

Abhängigkeit des ersten und zweiten Formanten veranschaulicht. Im zwölften Monat (Abb. 2) steht die Lallperiode in voller Entwicklung, die Klänge der Kinderstimme füllen diffus das ganze Gebiet aus, welches dem Vokalgebiet der Erwachsenen sehr ähnlich ist, nur ist es zu höheren Frequenzen verschoben. Die Realisation der Laute ist hier aber mehr oder weniger zufällig, so daß deren Differenzierung nur durch das Gehör des Untersuchers durchgeführt werden kann.

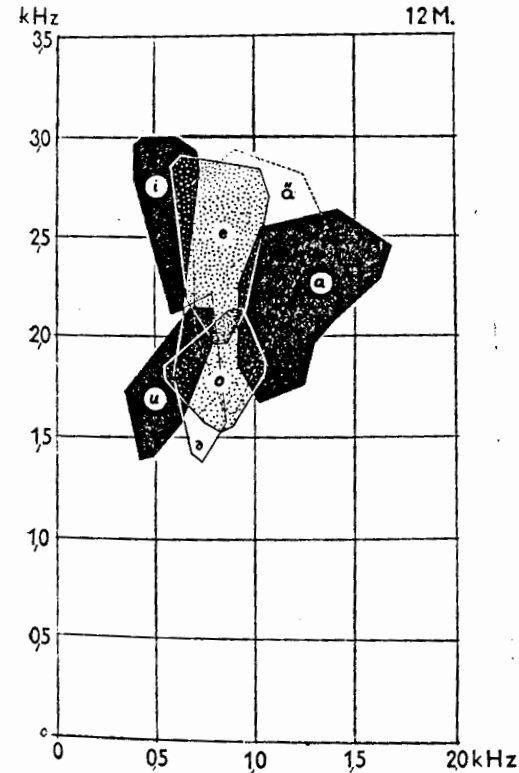


Abb. 2.

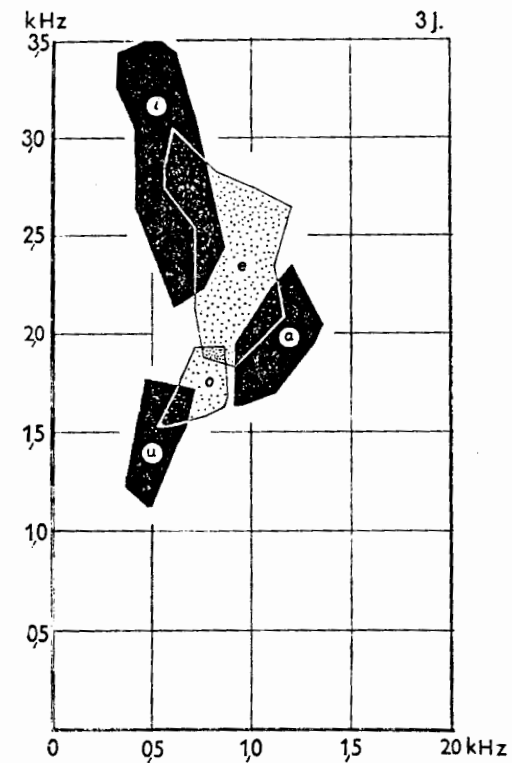


Abb. 3.

In der Nachahmungsperiode von Worten die sich bei den meisten Kindern im 15. Monat einstellt, sind die Vokale schon besser differenziert.

Im Stadium der ersten selbständigen Worte im 18. Monat sondern sich die Formantengebiete der einzelnen Vokale noch mehr ab. Besonders entfernen sich die Formanten der Selbstlaute *i* und *u* voneinander, was zu ihrer klanglichen Differenzierung wesentlich beiträgt.

Bei zweijährigen Kindern, die schon begonnen haben in Sätzen zu sprechen, schreitet die Differenzierung im gleichen Sinne weiter fort. Einen großen Ausmaß hat aber noch immer die Area des Lautes *a*, dessen Aussprache bei tschechischen Kindern sehr schwankend und breit ist.

Das letzte Bild (Abb. 3) zeigt die Formantengebiete dreijähriger Kinder. Sie sind

bereits gut differenziert und gut entwickelt, aber die ganze Area der kindlichen Selbstlaute bleibt ständig in einer wesentlich höheren Lage als bei den Erwachsenen und zwar wie im ersten so auch im zweiten Formanten. Aus dem Diagramm ist es möglich, besondere Charaktere der Färbung von tschechischen Lauten zu ersehen, worin ich mich jedoch im heutigen kurzen Referat nicht einlassen kann.

Zusammenfassend möchte ich noch konstatieren, daß sich die spezifische Färbung der Säuglingsstimme, die wir durch ihre Formantengebiete definiert haben, unter dem Einfluß der Hörwahrnehmung und der Ausreifung des Nervensystems fortlaufend ändert. Dies geschieht zuerst durch eine diffuse Vermehrung von Lauten, die sich später unter dem Einfluß des umgehenden Sprachmilieus in Lautstereotypen organisieren. Obzwar sich die Formanten der Kinderstimme immer mehr den Gebieten der Erwachsenen nähern, bewahren sie stets ihre höhere Lage, die durch die anatomischen und physiologischen Verhältnisse der Resonanzhöhlen der Sprachwerkzeuge und der Größe des kindlichen Kehlkopfes gegeben ist.

DISCUSSION

Slama—Cazacu:

Les deux communications (Mme Sedláčková et M. Truby) touchent un problème qui commence à être exploré avec des procédés objectifs. Des recherches de ce genre sont utiles non seulement pour le problème du développement du langage, mais aussi pour la phonétique générale. Elles m'ont intéressée aussi d'un point de vue personnel: nous avons initié une recherche, en collaboration internationale, concernant la formation du système phonématique chez l'enfant et des recherches de détail, si minutieuses, comme celles de Mme Sedláčková, nous seront très utiles. Je voudrais souligner aussi la nécessité de réaliser — par une coopération — un système international de transcription phonétique adéquat au pré-language enfantin.

Handzel:

Ich wollte die Autorin fragen, ob chronische Krankheiten [des Kehlkopfes (z.B. Dysphonien), Nasenhöhle usw.] auf die Formierung der Formanten irgendwelchen Einfluß haben (Observation betrifft einen 3-jährigen Zeitraum).

Sovijärvi:

Können die von Ihnen untersuchten 3-Jährigen Kinder die phonematische Opposition [i/e] so realisieren, daß man fast nie unsicher sein kann, was die Kinder meinen? In dem Fall, daß die Realisationsfähigkeit sich schon so weit entwickelt hätte, könnte man vielleicht die Frequenzen des 3. Formanten statt des 2. für das betreffende Diagramm der akustischen Vokalgebiete auswählen, weil der 3. Formant in dem recht gut differenzierenden Sprechen für das [i] entscheidend ist.

Witloch:

Différents points de vue possibles: 1) p. de vue strictement *phonétique*, 2) p. de vue *phonologique*, 3) faire la différence entre le degré de perfection atteint chez *un individu* et le degré de perfection pour ainsi dire „typisé“ (c' est celui de Mme Sedláčková, je crois). On peut „perfectionner“... presque à l'infini“... (Différents degrés de perf.: milieu familial, école, récitation, etc. etc.) Mme Sedláčková croit-elle qu'on pourrait fixer (trouver) un certain âge, quant à cette perfection?

AN ACOUSTICAL SIGNAL IN MANDARIN (CHINESE) QUESTIONS WITH INTERROGATIVES*

YAO SHEN**

There is a set of "bound words" and "free words" (Y. R. Chao, *Mandarin Primer*, Cambridge, 1948, p. 33) in Mandarin which we will call interrogatives. Some of the words are *shéi* (who/m), *shémma* (what), *něi-* (which), and *dzěmma* (how). These words have simultaneous semantic and grammatical functions. Each of these interrogatives distinguishes itself from the rest of the set by the lexical content it elicits. The response to an interrogative gives the lexical information requested. Grammatically the position of an interrogative in a supplement question and of its corresponding answer is the same.

If Mandarin does not seem to use word order to indicate interrogation in supplement questions, then does it do so by intonation?

Two factors determined the selection of materials used in the present effort to answer this question. First, questions with interrogatives should be those produced in conversations rather than those produced in isolation. The conversations used here were taken from Y. R. Chao's *Mandarin Primer*. Galaxy magnetic recordings of tapes of the conversations were made with Ampex 354 at 7 1/2 speed from the F. P. 80002 Folkways Records accompanying the text. Second, since significant intonation variations usually occur at the end of utterances, the data collected centered around interrogatives in final position.

Occurrences of interrogatives with tone-2 followed by tone-0, 20, were selected for two reasons. First, there were many occurrences of *shémma*, which has the tone pattern of 20. Second, it might be that since the *ma* in *shémma* has tone-0, the behavior of the harmonics of *ma* could be interpreted as that of intonation.

There were three kinds of final contours of the 10th harmonic observed: 1. a continuous rise; 2. a level after the rise; 3. a fall after the rise.

An examination of the context in which these three kinds of final contours occurred with questions with interrogatives showed five relations. (A) Repetition questions ended in a rise. (B) Repetition questions ended on a level. (C) Repetition questions

* Grateful acknowledgements go to Prof. Ruth Crymes and Prof. Margaret Spencer of Univ. of Hawaii and Prof. Don G. Stuart of Georgetown Univ., Washington D. C.

** University of Hawai & Summer School, Georgetown University.

ended in a fall. (D) Supplement questions ended on a level. (E) Supplement questions ended in a fall. (F) Previous research shows statements end in a fall. (Yao Shen, *English Phonetics*, Ann Arbor, Michigan, 1966, p. 254.).

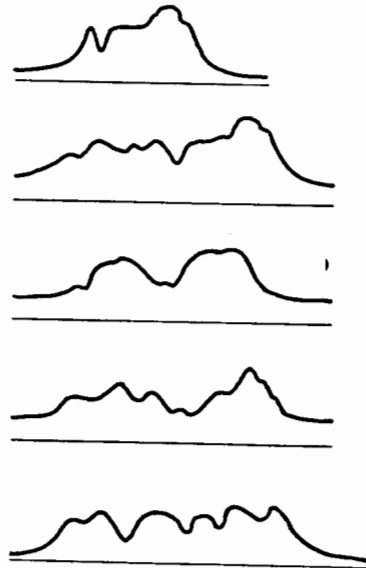


Fig. 1.

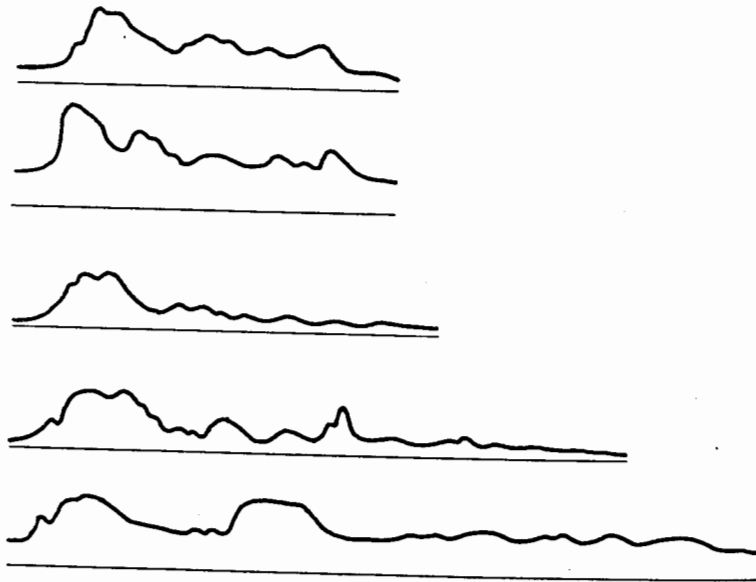


Fig. 2.

Interpretations of the data were made as follows.

One, statements (F) have the same final contour as questions with interrogatives (C) and (E). The request for a response distinguishes (C) and (E) from statements (F)

which do not request a response. The signal is lexical. Two, questions with interrogatives ending in a rise contour (A) are repetition questions. The signal is acoustical. Three, repetition questions and supplement questions have the same final contours. Examples are (B) vs (D) and (C) vs (E). What then is signalling the difference, since it is neither the lexicon nor the word order nor the final contour?

Patterns of amplitude-vs-time curves in the two kinds of questions seemed to show a reasonably consistent contrast.

First, in the repetition questions, the highest amplitude display coincides with the occurrence of the interrogative (Figure 1). Second, in the supplement questions, the highest amplitude display does not coincide with the interrogatives. It occurs somewhere else in the question (Figure 2). A tentative conclusion from this very limited data suggests an acoustical signal differentiating the two kinds of Mandarin questions with interrogatives. It is where the highest amplitude display is in the question.

DISCUSSION

Marquardt:

The data presented regarding Chinese questions is interesting. What are the implications of this data for the teaching of English to speakers of Mandarin Chinese or Chinese to speakers of English?

Shen:

ad Marquardt: Any detailed analysis which specifically describes a certain feature in a language is useful in language teaching. In this case, if this acoustical feature in Mandarin differs from what is used in similar English questions, when native speakers of Mandarin bring it into English, their English will sound Mandarin English. It also applies when native speakers of English learning Mandarin do not produce this feature, their Mandarin will sound non-Mandarin.

PHONTRNS: AN ALGORITHM FOR THE TRANSCRIPTION OF FRENCH TEXT INTO PHONETIC SYMBOLS BY A COMPUTER

GEORGETTE SILVA — BRUCE PRATT*

The purpose of this paper is to report a computer method for transcribing orthographic French into phonetic symbols. The procedure was developed at Monash University, Australia. Although similar schemes have been reported for English¹ and Swedish,² and work is in progress at Bonn on a system for transcription of German,³ as far as we are aware none has yet been published for French.

The system outlined in this paper is designed to accept text punched on paper tape in "running text" form. Where the computer's internal code does not include letters with orthographic signs such as accents, substitute symbols must be introduced at the typing stage. Other than this no precoding is required. The contents of the paper tapes are transferred to magnetic tape, proofread and corrected. The text is then ready for transcription.

Automatic transcription of modern French involves the resolution of a two-fold problem: on the one hand, certain sounds can be represented by a number of spellings; on the other, the same letter or group of letters is sounded in different ways in different words: in each specific case the sound must be defined in context in such a way that it is accurately identified.

This problem is greatly simplified if the text is first divided into syllables. This is done by computer, syllables being separated by blank spaces and words by asterisks, which provide convenient reference points during the process of transcription.

Since the computer cannot recognize phonetic symbols each sound is represented by a two digit number code: further, certain pairs of consonants have been allotted separate numbers to facilitate reference to them during processing.

Some combinations of letters regularly generate complex sound patterns and many difficulties can be eliminated by coding these before transcribing the remainder of the text. These groups have been arranged in lists according to the number of letters they contain, each item being accompanied by its pronunciation expressed in the number code referred to a moment ago.

* Monash University Clayton, Victoria, Australia.

¹ Reported by Francis F. Lee of M.I.T.

² Reported by Siv Engström.

³ At the Institut für Phonetik und Kommunikationsforschung, Bonn University.

The syllabized text is scanned on successive cycles in groups of four, three, two and one characters. Each time a group is formed it is compared with the list of like length. If a match is found, the numeric code is substituted for the group of letters found in the input text. On completion of each cycle an increasing proportion of the text has been encoded. This system provides for those groups of letters which never vary in pronunciation, and also permits some discrimination between different pronunciations of a given letter combination through use of precoding on early cycles. However this is not enough in itself. To ensure adequate transcription of those letters and groups of letters which vary in pronunciation from word to word, the basic procedure has been overlaid with an elaborate system of checks of adjacent letters, which in effect define the pronunciation of these groups in terms of adjacent sounds. Alternative codings are provided in a set of lists parallel to the main lists.

Our aim has been to develop a simple and flexible system for machine transcription, our concern being with the observance rather than the formulation of phonetic rules. On each successive cycle the problems facing the programmer can be stated afresh, as a certain number of possibilities have been eliminated: the only criterion used in deciding whether or not to accept a procedure has been the degree of accuracy it allows.

The real test of the system is therefore its accuracy. We are happy to report fewer than 50 errors on a text of more than 14,000 sounds, that is, an error rate of 0.4%, and we offer suggestions for rapidly locating and correcting these.

Output can be provided in any form desired, and some examples can be found in our recently published book.⁴ Some institutions may be hampered by the absence of print chains carrying phonetic symbols, but this would not greatly restrict the system's usefulness, as all significant operations are executed using the number code. It would seem that all that part of the phonetician's work which involves the statistical analysis of the language, and which up to the present time has had to be carried out manually by teams of research assistants on relatively small samples of the language, can now be performed on vast quantities of data. Studies that come to mind are the relative frequency of segmental phonemes, word length as measured by phonemes and syllables, canonical forms of consonant-vowel sequences arranged by frequency, consonant cluster distribution, transitional probabilities of phoneme sequences, entropy and redundancy values and so on.

In conclusion, we wish to emphasize that our aim has been to place in the hands of phoneticians a tool which will greatly increase the possible scope of their investigations. It is a highly flexible tool, open to considerable refinement and to modification according to particular modes of defining phonemes. We hope you will find it useful.

⁴ Bruce Pratt and Georgette Silva. *Phontrns: a Procedure which uses a computer for transcribing French text into phonetic symbols*. Monash University. 1967.

DISCUSSION

Valdman:

The program described here is based on the fallacy that French spelling is a sort of phonological transcription. In fact it is a morphonemic notation which provides information about form and substance. Therefore if we wish to devise a transcribing device whose output includes spontaneous speech, we shall need to provide it with higher level information i.e. information about syntactic structure and morphological form.

Gsell:

J'ajouterais peu de mots à ce que vient de dire Mr. Valdman, puisque le système graphique français n'est ni phonologique, ni phonétique, ni même morphologique. Je demanderais seulement à Madame Silva:

1. quel est le pourcentage d'efficacité de son système
2. s'il a été testé également sur la transcription de la prose.

Ayant fait travailler des chercheurs à Grenoble sur la transcription automatique, j'avais estimé que la structure du français écrit était tellement différente de celle du français oral qu'il fallait traiter les deux formes de l'expression comme 2 langues différentes et réaliser un programme de type traduction automatique avec morphologie du français oral et programme d'analyse syntaxique.

Silva:

ad Gsell: The system was tested on 'Le Jeune Parque' by Paul Valéry, a poem of some 500 lines. It is with some confidence that we report an accuracy of over 99%. The version of the program suitable for the transcription of prose is not yet completed.

INTERRELATIONSHIP OF ENGLISH CONSONANTS *

ADANAND SINGH**

A confusion-predicting model, based on distinctive feature relationships of English consonants, was proposed. Comparisons of the three different sets of classifications of interconsonantal relationships, offered by Miller and Nicely,¹ extended by Singh and Black² (MN ext. SB), by Halle³ (H), and by Wickelgren³ (W) were made to investigate which of these three systems approximated more closely the phonemic realization of English consonants. Errors in each distinctive feature category and combination of categories were compared to test the independence of features from the phonemic contexts. The distinctive feature system used for further analysis of data was MN ext. SB. Arbitrarily selected features were voicing, frication, duration, liquid, glide, retroflex, and place of articulation. Presence of a feature, for example, voicing, in the model was denoted by V^1 and its absence by V^0 . Similar notations followed for the other features except for place of articulation. Four places of articulation were chosen to characterize 22 consonants of English denoted by P^1 , bilabial; P^2 , alveolar; P^3 , palatal; and P^4 , velar. These were coded into a multi-category channel rather than into four independent channels.⁴

The phoneme (s) may be characterized by $V^0F^1D^1L^0G^0R^0P^2$, indicating that voicing, liquid, glide, and retroflex are not present while frication and duration are present at the alveolar place of articulation; the phoneme (g) may be characterized by $V^1F^0D^0L^0G^0R^0P^4$, indicating that frication, duration, liquid, glide, and retroflex are not present while voicing is present at the velar place of articulation.

* This study was supported by a grant from the National Institutes of Health (NBO-3388).

** Howard University Washington DC.

¹ G. A. Miller and P. E. Nicely, "An Analysis of Perceptual Confusion among some English Consonants," *J. Acoust. Soc. Am.* 27, 338—352 (1955).

² Sadanand Singh and J. W. Black, "Study of Twenty-Six Intervocalic Consonants as Spoken and Recognized by Four Language Groups," *J. Acoust. Soc. Am.* 39, 372—387 (1966).

³ Wayne A. Wickelgren, "Distinctive Features and Errors in Short-Term Memory for English Consonants," *J. Acoust. Soc. Am.* 39, 388—398 (1966).

⁴ Sadanand Singh, "Crosslanguage Study of Perceptual Confusion of Plosive Phonemes in Two Conditions of Distortion," *J. Acoust. Soc. Am.* 40, 635—656 (1966).

Table of Quotients of Distinctive Features of English Consonants.

	k	t	p	j	s	θ	f	h	d ₃	g	d	b	z	δ	v	l	r	j	w	tj	▲
k	U	p ²	p ³	F ⁻¹ p _D ⁻¹	F ⁻¹ p _D ²	F ⁻¹ p ₂	F ⁻¹ p ₃	F ⁻¹	V ⁻¹ F ⁻¹ p	V ⁻¹	V ⁻¹ p ²	V ⁻¹ p ³	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ²	V ⁻¹ F ⁻¹ p ₃	V ⁻¹ p ₂ L ⁻¹	V ⁻¹ p ₂ R ⁻¹ L ⁻¹	V ⁻¹ p ₃ G ⁻¹	V ⁻¹ p ₃ G ⁻¹	F ⁻¹ p	F ⁻¹ G ⁻¹ p ₃
t		U	p	F ⁻¹ p _D ⁻¹	F ⁻¹ D ⁻¹	F ⁻¹	F ⁻¹ p	F ⁻¹ p ⁻²	V ⁻¹ F ⁻¹ p ⁻¹	V ⁻¹ p ⁻²	V ⁻¹	V ⁻¹ p	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ²	V ⁻¹ F ⁻¹ p ₃	V ⁻¹ L ⁻¹	V ⁻¹ R ⁻¹ L ⁻¹	V ⁻¹ p ₃ G ⁻¹	V ⁻¹ p ₃ G ⁻¹	F ⁻¹ p ⁻¹	F ⁻¹ G ⁻¹ p
p			U	F ⁻¹ p _D ⁻²	F ⁻¹ D ⁻¹ p ⁻¹	F ⁻¹ p ⁻¹	F ⁻¹	F ⁻¹ p ⁻³	V ⁻¹ F ⁻¹ p ⁻²	V ⁻¹ p ⁻³	V ⁻¹ p ⁻¹	V ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻²	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p ₃	V ⁻¹ p ⁻¹ L ⁻¹	V ⁻¹ p ⁻¹ R ⁻¹ L ⁻¹	V ⁻¹ p ⁻² G ⁻¹	V ⁻¹ p ⁻² G ⁻¹	F ⁻¹ p ⁻²	F ⁻¹ G ⁻¹
f				U	p	p _D	p _D ²	p ⁻¹ D	V ⁻¹ D	V ⁻¹ F ⁻¹ p ⁻¹ D	V ⁻¹ F ⁻¹ p _D	V ⁻¹	V ⁻¹ F ⁻¹ p _D ²	V ⁻¹	V ⁻¹ p _D	V ⁻¹ p _D L ⁻¹	V ⁻¹ p _D R ⁻¹ L ⁻¹	V ⁻¹ F ⁻¹ G ⁻¹ D	V ⁻¹ G ⁻¹ F ⁻¹ p _D ²	D	DG ⁻¹ p ²
s					U	D	p _D	p ⁻² D	V ⁻¹ p ⁻¹ D	V ⁻¹ F ⁻¹ p ⁻² D	V ⁻¹ F ⁻¹ D	V ⁻¹ F ⁻¹ p _D ²	V ⁻¹	V ⁻¹	V ⁻¹ p _D	V ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ R ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ F ⁻¹ G ⁻¹ D	V ⁻¹ G ⁻¹ F ⁻¹ p _D ²	D	DG ⁻¹ p
θ						U	p	p ⁻²	V ⁻¹ p ⁻¹	V ⁻¹ F ⁻¹ p ⁻²	V ⁻¹ F ⁻¹	V ⁻¹ F ⁻¹ p _D ²	V ⁻¹	V ⁻¹	V ⁻¹ p	V ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ R ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ F ⁻¹ G ⁻¹ D	V ⁻¹ G ⁻¹ F ⁻¹ p _D ²	p ⁻¹ D	DG ⁻¹ p
f							U	p ⁻³	V ⁻¹ p ⁻²	V ⁻¹ F ⁻¹ p ⁻³	V ⁻¹ F ⁻¹ p ⁻¹	V ⁻¹ F ⁻¹ p _D ²	V ⁻¹	V ⁻¹	V ⁻¹ p	V ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ R ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ F ⁻¹ G ⁻¹ D	V ⁻¹ G ⁻¹ F ⁻¹ p _D ²	p ⁻¹	G ⁻¹ p
h								U	V ⁻¹ p	V ⁻¹ F ⁻¹	V ⁻¹ F ⁻¹ p ²	V ⁻¹ F ⁻¹ p ³	V ⁻¹ p _D ⁻¹	V ⁻¹	V ⁻¹	V ⁻¹ L ⁻¹ p ⁻¹ F	V ⁻¹ R ⁻¹ L ⁻¹ p ⁻¹ F	V ⁻¹ F ⁻¹ p ⁻² G ⁻¹	V ⁻¹ G ⁻¹ F	p ⁻²	G ⁻¹
d ₃									U	p ⁻¹ F	FP	FP ²	D ⁻¹			V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ₃
g										U	p ²	p ³	PF ⁻¹ D ⁻¹			V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
d											U	p	p ⁻¹ F ⁻¹ D ⁻¹			V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
b												U	p ⁻² F ⁻¹ D ⁻¹			V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
z													U			V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
δ																V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
v																V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
l																V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
r																V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
j																V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
w																V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
tj																V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²
▲																V ⁻¹ p _D ⁻¹	V ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p _D ⁻¹	p	G ⁻¹ p ²

V=voicing, F=fricative, P=place, D=duration, R=retroflex, L=liquid, G=glides, U=unity.

$$\text{Now the phonemes } (s) : (g) = \frac{V^0 F^1 D^1 L^0 G^0 R^0 P^2}{V^1 F^0 D^0 L^0 G^0 R^0 P^4} = V^{-1} F D P^{-2}$$

The inverse is $V F^{-1} D^{-1} P^2$. Thus, the ratio contains four features including an extra loading on the place category. Table 1 describes all the interconsonantal relationships of English in similar fashion. This is based on the MN ext. SB system of classification. One novel feature in this table is that once the first row has been obtained, and the relationship of (k) to the rest of the phonemes in the row has been computed, the subsequent relationships can be generated by following algebraic multiplication. For example, in order to find the relationship between (t) and (p) denoted by (t, p) the following relations are used: $(k, p) = (k, t)(t, p)$. Thus $(t, p) = (k, t)^{-1}(k, p) = p^{-2}p^3 = p$.

Native speakers of Hindi and English, 22 in each group, comprised the experimental subject. They recorded 22 pre- and 22 post-vocalic English consonants. Twentytwo native auditors each of Hindi and English responded to them in five S/N ratio (-8, -4, 0, +4, +8 dB) and five signal level (35, 30, 25, 20, 15 dB) conditions. Although the stimulus was never included in the response-choice, all remaining

V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p ₂	V ⁻¹ F ⁻¹ p ₃	V ⁻¹ p ₂ L ⁻¹	V ⁻¹ p ₂ R ⁻¹ L ⁻¹	V ⁻¹ p ₃ G ⁻¹	V ⁻¹ p ₃ G ⁻¹	F ⁻¹ p	F ⁻¹ G ⁻¹ p ₃
V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹	V ⁻¹ F ⁻¹ p	V ⁻¹ L ⁻¹	V ⁻¹ R ⁻¹ L ⁻¹	V ⁻¹ p ₃ G ⁻¹	V ⁻¹ p ₃ G ⁻¹	F ⁻¹ p ⁻¹	F ⁻¹ G ⁻¹ p
V ⁻¹ F ⁻¹ p _D ⁻¹	V ⁻¹ F ⁻¹ p ⁻¹	V ⁻¹ F ⁻¹	V ⁻¹ p ⁻¹ L ⁻¹	V ⁻¹ p ⁻¹ R ⁻¹ L ⁻¹	V ⁻¹ p ⁻² G ⁻¹	V ⁻¹ p ⁻² G ⁻¹	F ⁻¹ p ⁻²	F ⁻¹ G ⁻¹
V ⁻¹ p	V ⁻¹ p _D	V ⁻¹ p _D ²	V ⁻¹ p _D L ⁻¹	V ⁻¹ p _D R ⁻¹ L ⁻¹	V ⁻¹ F ⁻¹ G ⁻¹ D	V ⁻¹ G ⁻¹ F ⁻¹ p _D ²	D	DG ⁻¹ p ²
V ⁻¹	V ⁻¹ D	V ⁻¹ p _D	V ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ R ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ F ⁻¹ G ⁻¹ D	V ⁻¹ G ⁻¹ F ⁻¹ p _D ²	p ⁻¹ D	DG ⁻¹ p
V ⁻¹ D ⁻¹	V ⁻¹	V ⁻¹ p	V ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ R ⁻¹ L ⁻¹ D ⁻¹	V ⁻¹ F ⁻¹ G ⁻¹ D	V ⁻¹ G ⁻¹ F ⁻¹ p _D ²	p ⁻¹	G ⁻¹ p
V ⁻¹ p _D ⁻¹	V ⁻¹ p ⁻¹	V ⁻¹	V ⁻¹ L ⁻¹ p ⁻¹ F	V ⁻¹ R ⁻¹ L ⁻¹ p ⁻¹ F	V ⁻¹ F ⁻¹ p ⁻² G ⁻¹	V ⁻¹ G ⁻¹ F	p ⁻²	G ⁻¹
V ⁻¹ p _D ⁻¹	V ⁻¹ p ²	V ⁻¹ p ₃	V ⁻¹ L ⁻¹ p ² F	V ⁻¹ R ⁻¹ L ⁻¹ p ² F	V ⁻¹ F ⁻¹ p ₃ G ⁻¹	V ⁻¹ G ⁻¹ F ⁻¹ p ₃	p	G ⁻¹ p ₃
p _D ⁻¹	p	p ²	FPL ⁻¹	FPR ⁻¹ L ⁻¹	FG ⁻¹	FP ² G ⁻¹	V	G ⁻¹ p ²
p ₂ F ⁻¹ D ⁻¹	p ₂ F ⁻¹	p ₃ F ⁻¹	p ₂ L ⁻¹	p ₂ R ⁻¹ L ⁻¹	p ₃ G ⁻¹	p ₃ G ⁻¹	PF ⁻¹ V	F ⁻¹ G ⁻¹ p ₃ V
F ⁻¹ D ⁻¹	F ⁻¹	PF ⁻¹	L ⁻¹	R ⁻¹ L ⁻¹	p ₃ G ⁻¹	p ₃ G ⁻¹	p ⁻¹ F ⁻¹ V	F ⁻¹ G ⁻¹ p ₃ V
p ⁻¹ F ⁻¹ D ⁻¹	p ⁻¹ F ⁻¹	F ⁻¹	p ⁻¹ L ⁻¹	p ⁻¹ R ⁻¹ L ⁻¹	p ⁻² G ⁻¹	G ⁻¹	p ⁻² F ⁻¹ V	F ⁻¹ G ⁻¹ V
	p _D	p _D ²	PFDL ⁻¹	PFDR ⁻¹ L ⁻¹	FDG ⁻¹	p ² FDG ⁻¹	VD	DG ⁻¹ p ₂ V
U	D	p _D	FDL ⁻¹	FDR ⁻¹ L ⁻¹	p ⁻¹ FDG ⁻¹	PFDG ⁻¹	p ⁻¹ VD	DG ⁻¹ p ₂ V
	U	p	FL ⁻¹	FR ⁻¹ L ⁻¹	p ⁻¹ FG ⁻¹	PFG ⁻¹	p ⁻¹ V	G ⁻¹ p ₂ V
		U	p ⁻¹ FL ⁻¹	p ⁻¹ FR ⁻¹ L ⁻¹	p ⁻² FG ⁻¹	FG ⁻¹	p ⁻² V	G ⁻¹ V
			U	R ⁻¹	p ⁻¹ G ⁻¹ L	p ₃ G ⁻¹ L	p ⁻¹ L ⁻¹ V	F ⁻¹ G ⁻¹ L ₃ V
				U	p ⁻¹ G ⁻¹ L _R	p ₃ G ⁻¹ L _R	p ⁻¹ L ⁻¹ V _R	F ⁻¹ G ⁻¹ L ₃ V _R
					U	p ²	G ⁻¹ V	F ⁻¹ p ² V
						U	p ⁻² G ⁻¹ V	F ⁻¹ V
							U	G ⁻¹ p ²
								U

consonants (21) were paired with each other using ABX. All listeners did not hear all speakers; however, they heard all consonants in each condition of distortion.

The four confusion matrices relating to the prevocalic stimuli and four to the post-vocalic ones were analyzed for speakers and auditors. Rank-correlations were obtained between the eight confusion matrices and the three interconsonantal relationships. The responses to each of the 22 stimulus consonants were ranked from high to low and the number of distinctive feature differences from low to high on the assumption that with fewer differences between the two sounds, greater confusion occurs.

The error responses to the 22 pre- and 22 postvocalic consonants of English in four speaking-listening conditions were correlated with the three different systems of classifying interconsonantal relationships. To formulate a comparable response-predicting model bases on H and W systems, their respective notations were used.

Out of 528 rank-correlation values 270 were found significant either at the .05 or .01 level, *df.* 20. In pre-vocalic conditions, 157 of 264 were significantly correlated as compared to 113 in the post-vocalic condition. A greater number of significant rank correlations were found in the listening modes (158 of 264) than in the speaking modes

(112 of 264). The greater number of phonemes correlated with the model based on MN ext. SB system than on H system. The lowest number of significant rank correlations were obtained with W systems. The numbers were MN ext. SB 108, H 90, and W 72, each of 176 possible correlations.

The consonants that were commonly unpredictable in all the 4 speaking-listening conditions using MN ext. SB classification in prevocalic position were /z r/ and postvocalic /dlj/; using H system-prevocalic /hðlr/ and postvocalic /θhðlrj/; and using W system — prevocalic /hzj/ and postvocalic /bzvrhwjld/.

The predictability of the distinctive feature model based on MN ext. SB was tested further by comparing the scores in each of seven distinctive feature classifications and also in the feature combination categories of two's, three's, four's, and five's. The X^2 comparisons of the scores within a given category of feature showed no significance either in pre- or post-vocalic stimulus conditions. They were as follows: voicing, df 15; place, 17; frication, 9; duration, 7; retroflex, 1; liquid, 1; glide, 3; and combinations of two's df 109; three's 127; four's 105; and five's, 45 in both experiments. Thus, of a given feature or combination of features contrasting one pair of phonemes did not yield significantly different frequencies of errors as compared to contrasting another pair of phonemes.

DISCUSSION

Black:

I infer that your interesting procedure is based on an assumption that one distinctive feature equals another in aural effect. Hence the 50 per cent of significant correlations must be encouraging. But may not the remaining 50 per cent of non-significant correlations suggest further explorations of this basic assumption?

Singh:

Ad Black: It has been demonstrated earlier that all distinctive features do not preserve in errors equally. The results of the present study show that the error in a given distinctive feature category (inspite of the fact that they were from different phonemic contexts) did not yield significantly different results.

The fact that 50% of the consonants correlated with the model is self revealing that the present number of distinctive features is not adequate to characterize all consonants of English with significant perceptual relevance.

DIE TYPOLOGIE DER LAUTSYSTEME

V. SKALIČKA

Der Terminus Typologie ist mehr aus der Grammatik bekannt. Dort spricht man von Typologie und von Typen — z. B. vom agglutinierenden oder flektierenden Typus, vom Zusammenhang verschiedener Erscheinungen, von möglichen und unmöglichen Sprachen usw.

Im Rahmen der Lautlehre sind die typologischen Methoden verhältnismäßig neu. Man hat sich bisher vor allem mit dem gegenseitigen Verhältnis von Vokalen und Konsonanten oder mit der Kombinatorik beschäftigt. Viel wurde auch über Mikrotypologie der Lautlehre, d. h. über Typologie eng verwandter Sprachen geschrieben.

Ich möchte ein paar Worte über die Möglichkeit einer Typologie der Lautsysteme sagen.

Meinen Ausgangspunkt bilden Phoneme, die mehr oder minder Universalien sind. Diese Universalien geben ein klares Bild, etwa das folgende:

<i>i</i>	<i>u</i>	<i>p</i>	<i>b</i>	<i>v</i>	<i>m</i>	
<i>e</i>	<i>o</i>	<i>t</i>	<i>d</i>	<i>s</i>	<i>n</i>	<i>l</i> oder <i>r</i>
<i>a</i>		<i>k</i>	<i>g</i>			

Unklar ist in dieser Hinsicht die Rolle der Phoneme *h* und *j*.

Die Unterschiede der natürlichen Sprachen gegenüber diesem Schema sind von dreierlei Art: Einige Sprachen ersetzen die angeführten Phoneme durch andere, die mit ihnen wesensverwandt sind, z. B. deutsch *p* — *B* statt *p* — *b*, ungarisch *á* statt *a* usw. Andere Sprachen reduzieren die angeführte Liste: dem Finnischen fehlen *b*, *g*, dem Tahitischen *b*, *d*, *g*, *t*, *s* usw. Fast alle Sprachen vermehren die angeführte Liste um wenig oder mehr andere Phoneme (weitere Okklusivreihen, um Sibilanten, Affrikaten, lange Vokale und Konsonanten, nasalierte Vokalen). Ich möchte hier die Regeln dieser Vermehrung der Phonemzahl feststellen.

Es handelt sich hier um dreierlei Regeln: 1. Die Regel der Symmetrie. Die Lautsysteme neigen zur Symmetrie. Wenn sie ein *e* besitzen, so besitzen sie wahrscheinlich auch *o*. Wenn eine Sprache *p*, *t*, *k*, *b*, *d* besitzt, so besitzt sie wahrscheinlich auch *g*. Eine Sprache mit einem langen Vokal (oder Konsonant) besitzt mehrere lange Vokale (Konsonanten).

2. Die Regeln der Hierarchie. Manche Phoneme setzen einander voraus. Wenn z. B. eine Sprache eine Konsonantenmodifizierung besitzt (z. B. Mouillierung,

Aspiration oder Affrikation), so besitzt sie wahrscheinlich auch diese Modifizierung als ein selbständiges Phonem. Die Aspiraten setzen ein *h* voraus, die Palatalen ein *j*, die Assibilaten *c*, *č*, *č* ein *s*, *ś*, *š* (Ausnahme: spanisch *ch* ohne *š*).

3. Regeln der Inkompatibilität. Es gibt Beispiele dafür, daß gewisse phonologische Eigenschaften einander ausschließen, so z. B. Polytonie und Mouillierungskorrelation der Konsonanten.

Das sind die wichtigsten Einschränkungen der möglichen Sprachsysteme. Sonst sind die Möglichkeiten in hohem Maße frei. Die Sprachen entwickeln verschiedene akustisch-motorische Themen und kombinieren verschiedene Teile dieser Themen. So entwickelt das Russische und noch andere baltoslavischen und finnischugrischen Sprachen die Weichheit der Konsonanten (*p'*, *b'*, *t'*, *d'*, *n'* usw.). In allen Sprachen ist diese Erscheinung mit starker Entwicklung der Sibilanten verbunden (z. B. russisch *s*, *ś*, *š*, *z*, *ž*, *ž*, *š̄*, *ž̄*). Dagegen besitzen die kaukasischen Sprachen viele Sibilanten, aber keine Palatale. Für andere Sprachen ist die Konsonantenlänge das wichtigste Thema (Italienisch, Finnisch, Ungarisch), sie ist aber mit verschiedenen anderen Themen verbunden. Dasselbe gilt von der abnormalen Entwicklung der Spiranten (Avesta, Neugriechisch, Niederländisch). Dasselbe gilt auch von der Problematik der Vokale: die Vokale entwickeln in verschiedenen Sprachen Nasalität, Länge und suprasegmentale Einheiten (Akzent, Intonation).

Wir fassen zusammen: Die Typologie der Lautsysteme ist fast vollkommen selbständig. Das Wesen dieser Typologie liegt in der Entwicklung eines oder mehrerer akustischer Themen. Die Regeln, die hier herrschen, sind nicht in bestimmten Typen konzentriert.

DISCUSSION

Birnbaum:

Wenn sich keine phonologische Sprachtypen feststellen lassen, dürfen wir dann von einer Typologie der Lautsysteme sprechen? Vielleicht wäre es besser, einfach von verschiedenen phonologischen Sprachstrukturen zu sprechen. Sollen wir von der Thematik der Phoneme oder eher von derjenigen der (distinktiven) Merkmale sprechen? Einige der von Skalička angeführten Regeln der Inkompatibilität, z.B. Unvereinbarkeit der Polytonie und der Palatalitätskorrelation, erscheinen mir etwas fragwürdig (vgl. etwa das Serbokroatische).

Mańczak:

Die Regel 1 ist falsch. Obwohl Symmetrie manchmal in den sog. Lautsystemen vorkommt, ist wesentlich, daß in der Mehrheit der Fälle nicht Symmetrie, sondern Asymmetrie herrscht (es gibt z.B. kaum eine Sprache, wo die Anzahl der Palatalkonsonanten der der Nichtpalatale gleich wäre).

* Der Vortrag ist in *Phonetica Pragensia* (Acta Universitatis Carolinae — Philologica 6) Praha 1967, S.73f., unter dem Titel „Phonologische Typologie“ erschienen.

Vor Greenberg habe ich das Gesetz formuliert, demzufolge die Differenzierung nicht nur der Laute, sondern aller Sprachelemente von der Frequenz abhängt (z. B. der Ind. weist meistens mehr Tempora auf als der Imper.). Vgl. BSL 61, S. XXVIII ff.

Skalička:

Ad *Birnbaum*: Es handelt sich tatsächlich um eine Typologie ohne Typen. Einen besseren Namen weiß ich nicht. Das Beispiel von Inkompatibilität der Mouillierungskorrelation und der Polytonie haben die Phonologen gefunden. Ein anderes Beispiel wäre vielleicht richtiger.

Ad *Mańczak*: Es ist eine bekannte Tatsache, daß die merkmalthaltigen Phoneme zahlreicher sind als die merkmallosen. Das widerspricht nicht der angegebenen Regel 1.

THE ENGLISH DIPHTHONGS

ALENA SKALIČKOVÁ*

I am supposed to be speaking about "Diphthongs in English and in Czech", but as the time is so strictly limited, and as the situation concerning English is rather complicated I must leave the Czech material aside for today and try to have a cosmic-speed look at the English diphthongs only.

As you know these phenomena have been evaluated and their nature has been explained in several different ways. Theoretically there can be the question

1. of real diphthongs, i.e. of a connection of vowel + vowel,
2. of a diphthongoid, i.e. of a connection of vocalic element + a kind of consonantal element,
3. of so-called glide vowels (or vowel glides), and
4. of vocalic elements with modifications, or, in other words, of a palatalized, centralized or velarized ending of the representatives of the single tamber categories.

So let us consider these different possibilities.

(1)—(2): Now first what in fact is the difference between what is called a "diphthong" and what is called a "diphthongoid"? I do not see any. In each case one of the elements is vocalic (or prominent in the traditional terminology), which means that it is identified in its top-phase, i.e. in its steady-state phase, the second component being then identified in its on-glide phase only, i.e. in its transition from the preceding element before the so-called proper articulation or top-phase is reached, which fact makes these non-prominent components similar to the consonantal articulations. Therefore, I do not think it necessary to make any difference between diphthongs and diphthongoids, so that these two categories fall into one. —However, placing the phenomena in question under either of these headings is contradicted first by certain tamber differences of some of the initial elements from the nearest simple vowels, and secondly by the variations of the final components that are said to be usually "not reached".—Thus the first two categories do not seem quite adequate.

As for the third possibility, i.e. considering these phenomena as gliding sounds, this does not seem to comply with reality either, because if it really were the question of a constantly changing sound, how would it be possible to identify, e.g., its beginning as an [a]-sound differing either from [A] or from [a:]. Besides, the sonagrams show

* Institute of Phonetics, Charles University, Prague.

relatively long steady-states for the first elements. Obviously this "glide vowel" view is not fully tenable either.

However, the fact remains that these phenomena are different from the so-called simple vowels, that they stand in opposition to each other as well as to the other vocalic elements of English, and, further, that their beginnings differ from their ends. Besides, it is obvious that in these phenomena there is some process going on—be it a change in the character of their single components (as, e.g., in the case of [ou > əu]) or be it a question of monophthongization (as in [ɔə > ɔ:]) or a change into another sound complex or their merging (as in the original [uə] becoming [ɔə] that later gave [ɔ:], or in [au, əu] becoming [aɪ, əi] as Mr. Eustace says in his Congress paper). And at the moment it is not quite clear, where all these changes are going to lead, whether there will occur some common trend in all the phenomena concerned or whether the members of this group will part and either merge with the other vocalic categories of form an independent new one.

Now how to evaluate this complicated situation? At this transitory period in the development of these phenomena it is rather difficult to find a large enough category to fit them all in. Therefore, the fourth theoretical possibility mentioned at the beginning, i.e., a category of vowels with modifications (or modified vowels) seems, in the given circumstances, quite an acceptable temporary means of expressing the present status.

This stand-point could also explain the somewhat different tamber of the beginnings of the first elements from those of the nearest simple vowels as well as the "not reaching of the final elements", and, further, it could explain also the so-called monophthongization. As we know the opposition between the simple vowels is primarily based on tamber. With the so-called diphthongs it would be, therefore, rather suspicious that there occurs in their first elements a third functional tamber, and, moreover, that it may vary (because usually on the contrary only those phenomena that are not fully functional suffer remarkable variations in their acoustic realizations). And, further, just as only two degrees of quantity can usually reliably function within one language system while a third degree would have to be supported by some additional sound-feature, we can suppose that with the so-called diphthongs in English it could be a question of supporting a third tamber category (occurring within each of the tamber areas) by means of a modification. So that besides the two categories of the traditionally called "short" and "long" vowels there would be established a category of let us say "modified vowels", i.e., of vowels accompanied by palatalization (for the i-diphthongs) or centralization (for the centring diphthongs) or velarization (for the u-diphthongs) respectively.

Fantastic as it may sound at first still I think that at the present stage of development of these phenomena the category of "modified vowels" is not really out of place in view of the fact that there is obviously a process going on making their former status chaotic and that we still need either for the theoretical or for the practical purposes to have some large enough category to fit in all of them.

DISCUSSION

Vachek:

Dr Skaličková's very stimulating paper has the positive features of explaining Southern British diphthongs monophonemically and, besides, of stressing that the vocalic system of ModE phonemes is in a state of flux. It appears, however, that she does not give a definite answer to our question. First, the limit separating the long and modified vowels is rather doubtful — [i:] and [u:] very frequently appears as [Ii] and [Uu]. Second, there are not binary, but often ternary or even quaternary oppositions between the short and modified vowel phonemes (e.g. [ɔ] : [ou], [ɔi], [ɔə]; [ɛ] : [ei], [ɛə], etc.).

Skaličková:

I am very sorry that I cannot fully agree with Professor Vachek. As for my not giving "a definite answer to our question"—I do not think there can be one when "the vocalic system is in a state of flux". In my paper I said that "a category of modified vowels seems, in the given circumstances, quite an acceptable temporary means of expressing the present status" (cf. above).—As far as the doubtfulness of "the limit separating the long and the modified vowels" is concerned, I am afraid that it is just the [Ii], [Uu] phonemic interpretation that is doubtful, for the so-called "inclination to diphthongation" has always been considered unfunctional in the English vocalic system, so that there is no point in regarding it as a relevant decisive feature. Moreover, according to spectrograms it seems to be rather a question of the influence of a change in the pitch occurring within the given syllable than of an independent basic shift in the formant structure comparable with that of what is usually referred to as diphthongs.—As regards the [ɔ] : [ou], [ɔi], [ɔə] and [ɛ] : [ei], [ɛə], etc. oppositions I should like to say first of all that some of the sounds in question no longer exist in these forms, [ou] having become [əu], [ɔə] having practically disappeared, and the present situation of [ɛə] not being very clear¹); secondly, I do not see any reason why the plurality of the oppositions should be considered a drawback, namely, why a simple (or "unmodified") vowel could not stand in opposition both to a palatalized and to a centralized or to a velarized type.

¹ Cf., for example, A. C. Gimson, "An Introduction to the Pronunciation of English", p. 138.

QUELQUES PRINCIPES PSYCHOLOGIQUES ET PSYCHOLINGUISTIQUES DANS L'ÉTUDE DE LA PERCEPTION DE LA PAROLE

TATIANA SLAMA-CAZACU

1.1. Dans cette étape du développement des sciences, la nécessité d'une approche interdisciplinaire et expérimentale s'impose dans la phonétique également.

1.2. Par suite de l'accumulation des données expérimentales et des faits d'observations on ne peut plus envisager uniquement une phonétique *in abstracto*, détachée de la langue réalisée — donc de la parole —, de la situation réelle de communication, des deux pôles de cette dernière — donc de la réception aussi —; une phonétique qui ne tienne pas compte de l'interdépendance des processus articulatoires et perceptuels, du fait que l'émission est réglée par les mécanismes perceptuels proprioceptifs et par ceux concernant l'émission de l'interlocuteur. L'envisagement interdisciplinaire, à partir d'une perspective psychologique et psycholinguistique, devient obligatoire en phonétique. Car, par l'essence même des phénomènes en question, cette perspective est implicite dans l'étude d'un phénomène psychique — la perception — et d'un phénomène psycholinguistique — la réalisation de la langue, la parole.

2. Pour les recherches actuelles en phonétique, la méthode expérimentale est indispensable. Mais l'étude expérimentale phonétique de la perception de la parole ne peut point ignorer certains aspects importants de l'expérimentation psychologique de ce phénomène, ainsi que les résultats auxquels a abouti la recherche de la perception en psychologie.

2.1. Il nous faut souligner tout d'abord la nécessité d'expérimentations *de laboratoire* (où les conditions puissent être mieux contrôlées), et de l'utilisation d'*appareillages nouveaux*, exacts, qui enregistrent de manière précise les stimuli, la situation, les réactions et qui puissent, aussi, créer une simulation complexe, rapprochée le plus possible des conditions complexes de la réalité.

2.2. Un fait de majeure importance, qui semble un lieu commun, mais qui souvent est perdu de vue même en psychologie, est la *spécificité* (qui tient aussi au plan psychique et non seulement à la structure physique des stimuli) de la perception des stimuli verbaux, tant sur le plan de la perception auditive que sur le plan de la perception visuelle. Des expériences concernant la perception (visuelle pour le moment) des stimuli verbaux par comparaison avec les stimuli nonverbaux (P. Fraisse 1963,

1964, et nous même avec nos collaborateurs, 1966—1967)¹ — démontrent objectivement cette spécificité. En utilisant comme indicateur de la perception le temps de réaction soit verbale, soit motrice, nous avons établi (avec O. Vişan et M. Voicu) que les stimuli verbaux donnent un temps de réaction plus bas que les stimuli non-verbaux (ils donnent donc une perception plus rapide, grâce, très probablement, au fait que la verbalisation, donc le décodage — qui assure la perception complète d'un stimulus, en tant qu'élément d'une certaine catégorie de la réalité — est facilité par la structure même des stimuli verbaux offerts à la connaissance sous la forme du code-langue.

2.3. La perception des divers stimuli — donc des stimuli verbaux aussi — dépend de nombreux facteurs qui peuvent modifier la perception (et dont nous avons essayé d'établir² un tableau de 24 catégories, divisées en un grand nombre de sous-classes: modalité de structure des stimuli, état du sujet, facteurs motivationnels,³ technique expérimentale, conditions environnantes, complexité de la tâche perceptuelle, etc.). Ces facteurs multiples s'influencent réciproquement et varient eux mêmes en fonction des circonstances (donc de l'immixtion des autres facteurs). C'est ce que nous avons appelé le principe de la multiplicité et de la variabilité des facteurs de la perception, et qui devrait être pris en considération aussi en phonétique expérimentale, tout en essayant de montrer à quel degré et sous quelle forme concrète ces facteurs influencent la perception auditive du message.

2.4. Un problème difficile dans la recherche expérimentale de la perception est celui des indicateurs qui la révèlent. En phonétique expérimentale on s'est peu inquiété de l'importance de cet aspect méthodologique, qui peut falsifier bien de résultats. Si, d'une part, la psychologie doit encore chercher à établir la valeur objective

¹ P. Fraisse, Le temps de réaction verbale: dénomination et lecture. "Année psychol.", 1964, 64, nr. 1, pp. 21—46; T. Slama—Cazacu, O. Vişan, M. Voicu, The comparative efficiency of verbal and nonverbal stimuli (in condition of uncertainty and of verbal reaction time). "Rev. roum. sci. soc.-Psychologie", 1966, nr. 1, pp. 21—43; T. Slama—Cazacu, Use of motor reaction time for comparing the perception of verbal and non-verbal stimuli, and verbalisation in perception. "Rev. roum. sci. soc.-Psychologie", 1967, nr. 1, pp. 81—88; T. Slama-Cazacu, O. Vişan, Comparative study of perception of verbal and nonverbal stimuli, in complex conditions, by motor-reaction time, "Rev. roum. sci. soc.-Psychologie", 1966, nr. 1, pp. 65—79 (pour d'autres références bibliographiques nous nous permettons de renvoyer aussi à ces articles).

² T. Slama-Cazacu, Problèmes théoriques et méthodologiques de l'étude des facteurs de variabilité de la perception des stimuli verbaux et non verbaux et de leur efficacité, "Rev. roum. sci. soc. Psychologie", 1965, nr. 1, pp. 3—15; T. Slama—Cazacu, M. Voicu, Privire critică asupra studiului performanţei la diferiţi stimuli verbali şi neverbali, în psihologia experimentală contemporană a percepţiei, "Rev. psihol.", 1965, nr. 2, pp. 247—271.

³ R. Solomon, D. Howes, Word frequency, personal values and visual duration thresholds, "Psychol. rev.", 1951, 58, nr. 4, pp. 256—271.

des divers paramètres utilisés (réaction verbale ou déclaration du sujet, temps de réaction, réaction motrice ou végétative comme indice d'une perception etc.), d'autre part la phonétique expérimentale devrait élargir la sphère des indicateurs objectifs de la perception des stimuli verbaux et introduire dans ses moyens d'investigation certains de ces indicateurs déjà utilisés par la psychologie.

2.5. L'un des principaux obstacles pour le développement de la phonétique a été, probablement, l'envisagement statique des phénomènes sonores verbaux dans l'expérimentation mais aussi sur les autres plans de la recherche.

Nous voudrions insister ici sur la nécessité d'une étude dynamique de la perception de la parole.

a) Les stimuli verbaux, produit d'un être humain, sont destinés à être perçus par un être humain aussi. Si les recherches qui font appel à des modèles très abstraits, ou à des synthèses de sons, ou qui visent à détacher les traits essentiels qui caractérisent un son dans son émission ou qui le font perceptible comme tel dans la réception sont très utiles, on ne pourrait perdre de vue le fait que dans la communication il y a souvent une intervention de facteurs secondaires de l'ambiance, qui modifient l'émission d'un son ou sa réception.

Nous insistons donc sur la nécessité de placer les recherches phonétiques elles-aussi dans les ensembles réels, dans la communication (étudier la variation de la perception par rapport au déroulement de la chaîne articulaire, en fonction des divers ensembles, de la position contextuelle,⁴ de la longueur du message, des stimuli concomitants mimiques, gestuels etc.; le rapport entre la perception auditive du message verbal et la perception visuelle de ses composantes de motricité faciale reste encore à étudier et les résultats de ces recherches seraient très fructueux aussi pour la phonétique articulaire).

b) Un autre aspect de l'étude de la perception en tant que phénomène dynamique consiste à considérer celle-ci comme processus d'ajustement graduel, de prise de conscience par étapes des faits verbaux acoustiques et proprioceptifs. Le percept n'est pas le résultat immédiat et tout de suite correct d'un moment perceptif qu'on pourrait étudier en l'immobilisant par un instantané statique. Il y a dans la perception des moments qui se succèdent rapidement et au cours desquels le percept prend contour, se précise, est complété, est modifié même. La perception est d'ailleurs liée aussi à l'apprentissage, on „apprend“ à mieux discerner, à mieux percevoir. D'autre part, l'amélioration articulaire entraîne une amélioration perceptive, tout comme les données perceptives et proprioceptives règlent l'émission verbale. Le cercle perception-articulation-perception — que d'ailleurs on pourrait mieux dénommer progression cyclique — impose d'envisager dynamiquement la perception du langage.

⁴ P. Fraisse, Incertitude du stimulus et incertitude de la réponse dans la reconnaissance perceptive, "XVIIIth Int. Congr. psychol. Symposium: Information theory and perception". Moscou, 1966, pp. 23—46.

c) La connexion articulation — perception⁵ (ou vice versa) est l'un des traits d'une loi fondamentale du langage en général et de la réception en particulier.

L'étude phonétique ne peut être complète et correcte sans l'étude approfondie des particularités de la perception acoustique des sons verbaux. La perspective côté de l'émission enrichit l'étude phonétique.

On peut bien percevoir et ne pas être capable de prononcer correctement, mais ce qu'on n'a pas bien perçu, ne peut être correctement prononcé et ce qu'on prononce bien, a été et est bien perçu. L'émission correcte nécessite une réception antérieure correcte, mais la perception est aidée par une connaissance antérieure, par des jalons déjà tracés sur un schéma antérieur à la perception. Plus on vient à la rencontre du moment perceptif avec un schéma *compatible* (correspondant et non contradictoire à l'objet) plus complet, et plus la perception est facilitée, est correcte.

d) Enfin, l'étude dynamique de la perception des sons verbaux inclut aussi une étude génétique-évolutive de la perception, en fonction de la ligne du développement ontogénétique (mais aussi philogénétique).⁶ Le développement du langage et l'acquisition des langues étrangères sont basés sur l'apprentissage, donc sur une perception du langage. L'enfant entend et imite, il s'entend, il „se sent" parler et se corrige par un continuel feed-back. Le rôle de la médiation perceptive-auditive et proprioceptive articulaire est fondamental dans l'acquisition des langues.

DISCUSSION

Šabršula:

Je suis d'accord avec l'affirmation de Mme Cazacu que l'on ne peut pas détacher la phonétique de la parole. L'expression linguistique, quoique non motivée directement (*physei*, Cratyle) dépend du contenu, de la dynamique de l'énoncé. L'histoire des changements phonétiques prouve quelquefois que les éléments au service du rhème sont pertinents. Dans certains cas il est possible de prouver une dépendance *indirecte* qui va du contenu, de sa substance et de sa forme, jusqu'à la forme et la substance de l'expression.

⁵ Harlan Lane, The motor theory of speech perception. A critical review, "Psychol. review", 1965, 72, nr. 4, p. 292.

⁶ R. Chocholle, Relations entre audition et langage, „Rev. g-le des sciences", 1966, nr. 3—4, pp. 99—101.

WHAT CAUSES THE VOICED-VOICELESS DISTINCTION?

I. H. SLIS*

The difference between voiced (*v*) and voiceless (*vl*) consonants (*C*), in linguistics characterized by one distinctive feature, proves to be carried by a number of acoustical and perceptual attributes. They are: 1a) C_{vl} are longer than C_v . 1b) A vowel (*V*) adjoining a C_{vl} is shorter than one next to a C_v (with preceding [*V-*] more than with following [*-V*] vowels). Measurements of the duration of pairs of isolated Dutch words, differing in voice character only, yield no significant difference in mean duration; consequently it seems justified to assume that differences in duration of the *C* are compensated by differences in the *-V-* (fig. 1). 2) Both *V-* and *-V* seem to have longer

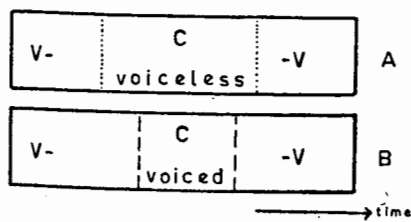


Fig. 1. Schematic build-up of VCV combinations with a voiceless [A] and a voiced [B] consonant.

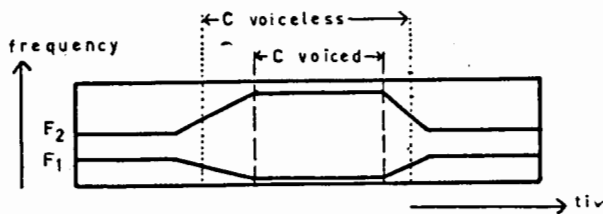


Fig. 2. Schematic combinations as in fig. 1 with addition of identical non-interrupted formant transitions.

formant transitions with C_v than those next to C_{vl} (duration as well as frequency range). By drawing identical formant transitions in fig. 1, we see that parts of the transitions, in the *v*-situation belonging to *V-*, correspond to the initial and final parts of *C* in the *vl*-situation (fig. 2). 3) During C_v a sound generated by vibration of the vocal cords is usually detectable; although, due to a damping of the higher frequencies, the F_2 will acoustically be absent, a low continuation of F_1 remains detectable (fig. 3B). During the C_{vl} this sound is absent (fig. 3A). 4) The sound level of the friction noise is higher in C_{vl} than in C_v . 5) The sound level of *-V-* next to C_{vl} is lower than that of those next to C_v (in *-V* more than in *V-*). 6) The F_0 contour of *-V* following C_{vl} starts high and declines during *-V*, whereas after C_v a rise of F_0 occurs in the initial

* Instituut voor Perceptie Onderzoek, Eindhoven.

part of -V followed by a fall after 50–100 ms. The top of the F_0 contour tends to higher values after and, to a lesser extent, before C_v than with C_{vl} .

Each of the attributes is perceptually detectable, although more attributes, rather than a single one, have to cooperate to obtain a change in voice character. To get a better understanding how these attributes are connected, an articulatory model will be developed. Differences in the physiological aspect of the v/vl -distinction as found in the literature are: 1. With C_{vl} , the larynx tends to a higher position than with C_v .

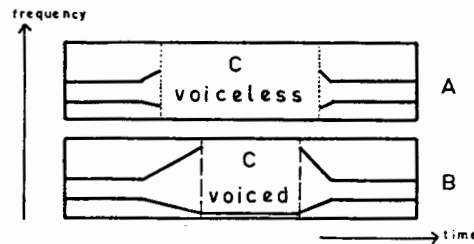


Fig. 3. Schematic build-up of VCV combinations as in fig. 1 and 2; the F_2 is interrupted during the consonants and F_1 is removed in the voiceless situation.

2. With C_v the pharynx seems to be wider than with C_{vl} . 3. The glottis is in vibration action during C_v while during C_{vl} it is slightly opened and not vibrating. 4. The intraoral pressure (P_{i0}) is higher with C_{vl} than with C_v . 5. With C_{vl} a stronger airflow occurs than with C_v .

The effect of a high larynx position and a narrow pharynx (C_{vl}), is a small volume of the pharynx; if the larynx is low and the pharynx wide (C_v) a large volume results. The resistance (R_{gl}) for the upward airstream is lower in the case of C_{vl} (open glottis) than in the case of C_v (vibrating glottis). The small volume and the low R_{gl} together can be held responsible for a quick build-up of the P_{i0} with C_{vl} ; contrary to this the large volume and high R_{gl} with C_v prevent a quick build-up, so that P_{i0} may not reach the value obtained with C_{vl} . Assuming that the suglottal-P is independent of the voice character of C, differences in P_{i0} will cause differences in P-drop across the glottis (increase of P_{i0} causes decrease of P-drop). The main reason of the cessation of voice with C_{vl} lies probably in too low a P-drop causing too weak an airstream to sustain vocal vibration.

As the vocal cords are attached to the larynx it seems probable that a difference in its position results in a difference in the condition of the vocal cords. This may explain the differences in F_0 and sound level within -V-; a high larynx position causes an unfavourable vibration condition of the vocal cords, manifesting itself in an abduction of the vocal cords. This condition results in a high F_0 and a low sound level (C_{vl}). The airflow depends directly on the P-drop across the oral constriction. A high P_{i0} leads to a high P-drop which causes a strong airstream (C_{vl}). This strong airstream accounts for the high sound level of C_{vl} .

Measurements of the total duration of word pairs, differing in voice character only,

yield no significant difference in mean duration. However, words with fricatives proved to be significantly longer than those with plosives and words with dentals proved to be longer than those with labials. These results show that if the mode of operation of muscles, controlling the open-close movement as well, changes, its consequences manifest themselves in a difference of duration. As a difference in voice character is not coupled to one in duration it seems justified to assume that the v/vl -distinction is controlled by muscles operating independently of the closing musculature. The pharyngeal constrictor muscle seems likely, since: 1. It does not interfere with the closing movement of the mouth. 2. It influences position and condition of the larynx. 3. In constituting the back wall of the pharynx, it will exert influence on the width of the pharynx.

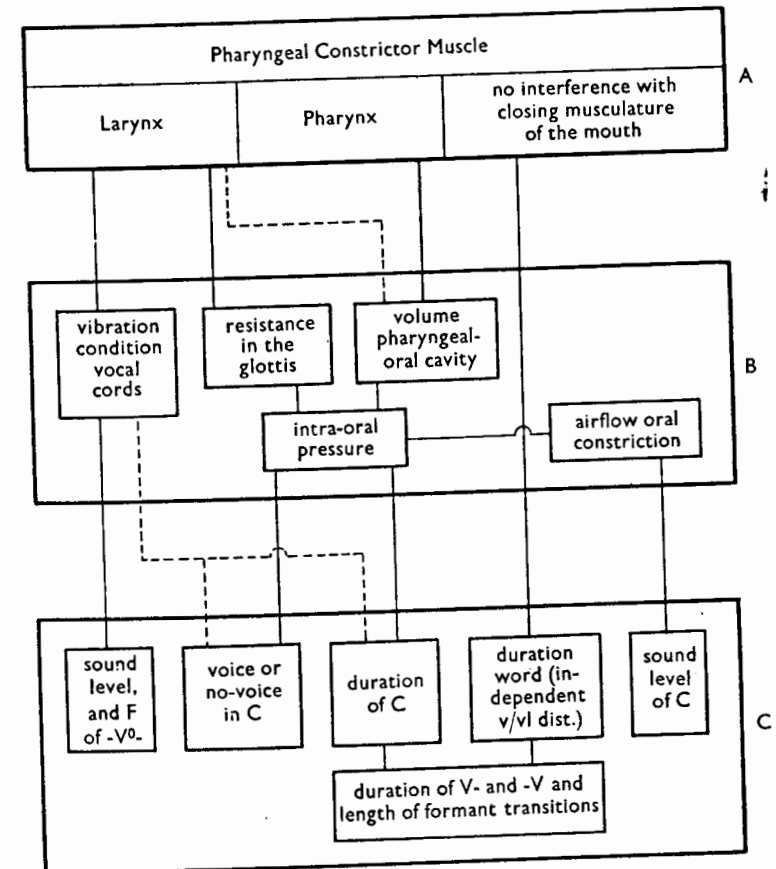


Fig. 4. Model accounting for the voiced-voiceless distinction, relating: the physiological factors [A], their physical effects [B] and the acoustical realisations [C].

These three physiological factors form the basis of the measurable attributes. A schematic view of these causal relations is represented in fig. 4 through the articulatory model.

DISCUSSION

MacCarthy:

Request for information on several points:

1. Were the categorical statements made by the speaker based on linguistic evidence from Dutch speakers?

2. How would the speaker deal with the fact that in different languages (e.g. English v. German, or Southern v. Northern English) there are different tendencies to shorten e.g. vowels before voiceless consonants? Would he say that the physiological tendencies are always operating, but that one language (dialect) gives way to or alternatively resists the tendency more than another?

Smith:

What is the background for your statement, that the larynx is raised by the activity of the constrictors and thus favours the unvoiced position?

When you only have this from literature I think, that X-ray film studies would be interesting. This may—or may not—support your statement.

Slis:

Ad MacCarthy: As an answer to your first question I can say that the acoustic attributes as mentioned are based on a series of isolated words spoken by Dutch speakers. The physiological attributes are partly based on material obtained with Dutch speakers and partly on material, found in the literature of speakers of other languages. The answer to the second question is: that it is quite possible that besides the actions described in the model other actions take place; e.g. that the command for a vowel preceding a voiceless consonant is different from one preceding a voiced one in certain dialects or languages.

Ad Smith: First of all I want to state that I made a model and that I do not pretend to make statements. Secondly that the part of the model under discussion is based on Literature only indeed.

LA PHONÉTIQUE ET L'AUDIOLOGIE

CARL-GUSTAF SÖDERBERG

C'est un fait bien connu que, sous certains rapports, les recherches phonétiques ont enrichi l'audiologie. Ainsi, l'analyse des transitions des formants a eu pour résultat, entre autres choses, une logopédie affinée à l'égard de ceux qui souffrent de graves troubles de l'audition. L'étude du pertinent dans la phonologie a exercé une influence importante sur le travail de l'audiologue qui, de plus en plus, se rend compte du rôle de la redondance. Enfin l'ensemble des considérations phonétiques et audiologiques ainsi que télétechniques et psychologiques a constitué la base audio-technique du laboratoire de langues.

Cependant, le phonéticien n'a pas encore utilisé l'audiométrie moderne de la manière qui soit à désirer. Les tests auditifs comprennent souvent des sujets dont on n'a pas vérifié l'état auditif par des audiogrammes. Une étude sur la perception de la parole devrait être regardée comme insuffisante si l'on risque d'y confondre les dates auditives avec les dates psychologiques. Bref, un test auditif aurait-il eu le même résultat si l'on avait compensé les défauts auditifs de certains sujets jusqu'au niveau normal?

En principe, il n'est possible d'appliquer le correctif pertinent dans l'enseignement quelconque, en particulier dans celui des langues, que si l'on a constaté l'état auditif de l'élève. Un audiogramme, dans les limites de la réponse du signal de la parole (70 Hz—7 à 10 kHz), devrait établir une base de méthode plus solide. Combien d'enseignants de langues vivantes, par exemple, savent leurs conditions auditives? Comment ces conditions influencent-elles leur enseignement? Quelle est la grandeur de la perte méthodique à cause de l'information peu satisfaisante sur l'audition du récepteur ou de l'émetteur?

On trouve aujourd'hui de l'équipement audiométrique qui peut être utilisé dans le laboratoire phonétique habituel (par exemple Kamplex Screening Audiometer Type TA 16). Il va sans dire qu'un audiomètre employé aux examens audiométriques rapides dans la routine d'un institut de phonétique aura une importance considérable. Si l'on installe un appareil du type indiqué ci-dessus la coopération entre le phonéticien et l'audiologue sera encore plus efficace.

Il y a des choses qui parlent en faveur d'un enseignement futur fondé sur la matière programmée émise par des systèmes de transmission de la parole. La motivation d'un développement pareil semble si puissante que l'on ne peut pas la négliger. Il

devrait être de toute importance que ce congrès prenne des mesures pour établir une norme internationale en ce qui concerne le laboratoire de langues, surtout du point de vue audio-technique. De telles dispositions exerceraient vraiment une influence fructueuse sur le développement de la pédagogie appliquée au laboratoire de langues.

Le phonéticien de nos jours s'intéresse certainement au progrès pédagogique. Comment le traitement de troubles de l'audition peut-il agir sur la pédagogie générale? Il est vraisemblable que l'audiologue professionnel puisse approfondir la méthode où se rencontrent des stimuli auditifs, visuels et tactiles, pour mettre le stimulus auditif au rapport avec d'autres stimuli. A ce sujet le rôle du psychologue sera manifeste.

Il y a en Suède, par exemple, des fonctionnaires chargés de la tâche de rechercher les élèves de l'école obligatoire ayant des troubles auditifs. Pour des raisons financières on ne poursuit pas ce travail à l'école supérieure. Il serait d'une grande valeur si l'on avait des renseignements audiométriques sur tout citoyen sans regarder l'âge. Depuis longtemps on considère les dates ophtalmiques correspondantes comme indispensables.

L'enseignement programmé paraît, selon des recherches récentes, indiquer de meilleurs résultats d'études que l'instruction traditionnelle en classe. Les conditions auditives d'un laboratoire de langues bien construit sont évidemment plus favorables que celles d'une salle d'enseignement habituelle. Voilà ce qui souligne l'importance d'améliorer la situation auditive dans les locaux conventionnels, entre autres par une réduction de la répercussion du son ainsi que par l'utilisation d'un type de laboratoire de langues simple (AP, AA). Après tout, la méthode orale nécessite des dispositions de cette espèce.

Il y a parfois des stagiaires souffrant de troubles de l'audition sans s'en rendre compte. Si l'on prend les mesures recommandées ci-dessus on éviterait des extrêmes pareils. Souvent c'est une question de troubles qu'on peut facilement écarter à l'aide d'un appareil de correction auditive. Cependant, l'acoustique inférieure dans la salle aggrave la situation auditive et évoque une certaine irritation chez l'enseignant ainsi que chez les auditeurs. En tout cas l'examen médical demandé par les autorités ne suffira pas au sujet des enseignants de langues. On devrait exiger un examen audiologique, qu'on exige déjà chez les enseignants spéciaux. Et il faudrait aussi élargir l'instruction de la technique de la parole.

EN RÉSUMANT JE VOUDRAIS BIEN SUGGÉRER:

1. Facilités audiométriques pour pouvoir faire des examens rapides chez les sujets des tests auditifs.
2. Norme audio-technique internationale pour les laboratoires de langues.
3. Examen audiologique obligatoire pour les enseignants de langues vivantes.

4. Coopération étendue entre les phonéticiens et les audiologues.
5. Mesures pour améliorer la situation acoustique dans l'enseignement.
6. L'élargissement de l'enseignement de la technique de la parole dans les études phonétiques.

DISCUSSION

Bystrzanowska:

Prof. Söderberg a touché un problème très essentiel. Je suis médecin, oto-rhino, et en écoutant certaines communications j'ai l'impression qu'il nous manque un langage commun. Plusieurs problèmes phonétiques dont on cherche la solution dans le domaine de psychologie, dans la phonologie, dans l'abstrait, pourraient être résolus facilement par les moyens audiologiques.

Florián:

L'audiologie est considérée comme une partie des épreuves ORL (otoscopie, examens classiques de l'ouïe) et ce n'est qu'après ces examens que doit suivre l'audiologie, qui fait partie du complexe.

De Grève:

1. J'applaudis à l'organisation de tests audiométriques organisés par le gouvernement suédois, et il serait utile: 1) que ces textes soient généralisés; 2) que les résultats en soient publiés. Il me semble également intéressant de limiter ces tests à des cas d'espèce et bien précis; j'ai moi-même fait faire des tests de ce genre dans 25 écoles secondaires sur chaque fois 5 élèves de valeur normale sauf en langue(s) étrangère(s): jusqu'à présent les résultats donnent 3,7/5 élèves souffrant de troubles audiologiques.

2. Je me réjouis de l'appel lancé par M. Söderberg pour une coopération entre phonéticiens et audiologues. J'aimerais également voir les pédagogues se joindre à cette coopération.

THE ACOUSTIC SIGNAL AND THE REFLEX THEORY

MILOŠ SOVÁK

The reflex theory based on Pavlov's teaching and amplified by the latest discoveries on neurophysiology, psychology and even by the theory of information leaves the concept of a closed reflex arc and accepts the thesis of an open reflex circuit. Stimuli from the environment, acoustic signals included, pass through the individual from the receptor to the effector and as a reaction return to the environment. On passing through the individual the stimuli are transformed, under pathological conditions deformed, even blocked.

In the receptive and central part certain phases of the reflex circuit play their role in the acceptance and transformation of acoustic signals, as follows:

1. Environment supplies acoustic signals. The acceptance of them by the individual depends on dialectic relations between the environment and the individual and depends on life constellation too: insignificant signals are not accepted as stimuli to reaction and extinguish gradually. Signals of both positive or negative importance i.e. signals of advantage or danger are accepted. It depends on the acoustic quality of signals too: too powerful or too frequent stimuli cause inhibition, too weak stimuli remain below the level of perception.

In no case is the acceptance of signals considered to be a mechanical passive process, on the contrary the activity of the individual plays its important role. It is known that the individual not only accepts but looks for stimuli — according to inner motivation caused by instinctive tension and by the necessities of life.

2. In the receptor the hearing and basic analysis of sounds is performed. The crackling sounds as phylogenetically older and complex sounds inclusive tones as phylogenetically younger are differentiated in Corti's organ.

3. The neural tracts and lower centres the reticular formation belongs to, have functional importance in transformation of signals through non-specific perception and therefore through affective accent. These functions which may be called listening, perform the lowest evaluation of signals by means of the orientation reflex.

4. The cortical area of the acoustic analyser analyses sound qualities, especially timbre. Evolutionally older regions of cortex (alocortex) take part in the evaluation of modulation factors; this may be considered the lowest degree of communication.

5. The conjunction of acoustic analyser with the other analysers performs the

basic (primitive) analysis and synthesis of signals. There imitative reflexes and conditioning take place.

On the level of the first signal system the psychological processes of recognition appear: i.e. perception, association, retention and recall of the complex ideas associated with acoustic signal. Through this gnostic function the signal is transformed into the information, the sound is transformed into the meaningful sound.

6. In the cortical areas of the dominant hemisphere the highest functions of human communication (on the level of the second signal system) are localized. We accept the three-phase division (according to the American authors—Wepman, Jones, Bock, Pelt):

a) in the input, recognition of the acoustic structure of complex verbal sounds takes place. It is essentially so called decoding (in the sense of the theory of information) or the same thing that Hardy calls “auding” and Lurija “phonematic hearing”.

The recognition—verbal gnostic function—is performed according to individual life experience.

b) In the integrating part, understanding of the contents of the acoustic information, i.e. the evaluation on the highest level, and programming takes place.

c) In the output corresponding reaction is formed through so called encoding (or verbal praxis).

7. In the expressive part of the reflex circuit the formed information is coordinated and realized as a signal to environment.

A further modification of the given information can occur according to the attitude of the acceptor of the signal.

We can see that both individual and environmental factors play an important role in transforming acoustic signals. The purpose of our contribution is to draw attention to the decisive role of feeling and to the individually variable activity of reflex processes through which acoustic signals are accepted and transformed.

ON TRANSITION IN THE LIGHT OF X-RAY FILMS

ANTTI SOVIJÄRVI

In my discourse at the IVth International Congress of Phonetic Sciences in Helsinki in 1961, I already touched to some extent the subject of transitions; namely transitions of the components of some Finnish diphthongs. For demonstrational purposes I used a special piece of equipment called ADAM (Apparatus for Demonstrating Articulatory Movements), which had just been developed at our Phonetic Institute. In the meantime some parts of this apparatus, especially the contourtape, plastic die plates, and the illumination showing through the pierced back plate have been improved to simplify the process of the diagramming of X-ray films and the photographing of the separate pictures thereof. Fig. 1.

Today I shall as examples handle some *VC-*, *VV-*, *CC-* and *CV-* transitions as they appear in some Finnish words and are shown in the successive single frames of X-ray films and in the corresponding segments of spectrograms.

When analysed transitions from that X-ray sound film for whose speed was chosen 48 frames/sec. I used the following method. Segments of speech corresponding to film frames were diagrammed with the help of the apparatus (ADAM) mentioned above. Various types of spectrograms were made of the series of words where the

Table 1. The transition of *pä* (in the word *täytyypä*)

		<i>ä 1</i>		<i>ä 2</i>		<i>ä 3</i>	
<i>F</i>	Print	<i>c/s</i>	<i>db</i>	<i>c/s</i>	<i>db</i>	<i>c/s</i>	<i>db</i>
<i>F0</i>	<i>NB</i>	321		318		311	
<i>F1</i>	<i>VP</i>	675	30	950	36	975	42
<i>F1</i>	<i>WB</i>	800		975		925	
<i>F1</i>	<i>NB</i>	642		954		933	
<i>F2</i>	<i>VP</i>	1950	36	1950	36	1900	36
<i>F2</i>	<i>WB</i>	1875		2025		2000	
<i>F2</i>	<i>NB</i>	1826		1908 + 2226		1866 + 2177	
<i>F3</i>	<i>VP</i>	2950	24	3200	24	3275	24
<i>F3</i>	<i>WB</i>	2925		3200		3275	
<i>F3</i>	<i>NB</i>	2889		3180		3110 + 3421	

transitions being examined occurred. For this purpose were used either the Kay Sonagraph (Model 661-A) or the Voiceprint Laboratories Sound Spectrograph (Model 4691 A) recently obtained by our Institute in Helsinki. For analization of the transitions the formant data were computed from the end section of the segments representing the frames. The actual speed of this film was 47,4 frames/sec.

In Fig. 2 is seen the normal spectrogram made with the Voiceprint Spectrograph using a wide band filter and flat shaping of amplitude. By vertical lines the divisions

Table 2. The transition of *tä* (in the word *täytyypä*)

		ä1		ä2		ä3		ä4	
F	Print	c/s	d/b	c/s	d/b	c/s	db	c/s	db
F0	NB	290		293		304		304	
F1	VP	285	12	900	30	950	36	925	30
	WB	700		900		925		925	
F2	NB	870		876		912		912	
	VP	2600	6	2375	24	2175	30	2100	24
	WB	2650		2375		2250		2125	
	NB	2610		2344		2128+2432		2128	
F3	VP	3700	18	3700	18	3675	18	3675	24
	WB	3675		3625		3650		3650	
	NB	3480 + 3770		3516		3648		3648	

NB = Narrow band, linear, high shaping

WB = Wide band, linear, high

VP = contour, high voiceprint spectrogram.

Table 3. The transition of *äy* (in the word *täytyypä*)

		ä8		ä9		ä10		ä11	
F	Print	c/s	d/b	c/s	db	c/s	db	c/s	db
F0	NB	370		385		397		416	
F1	VP	1100	42	950	30	825	36	825	42
	WB	1100		950		800		750	
	NB	1110		770+1155		794		832	
F2	VP	2400	36	2350	36	2350	42	2475	36
	WB	2450		2350		2175		2475	
	NB	2220+2590		2310		1985+2382		2496	
F3	VP	3575	30	3475	36	3550	36	3475	36
	WB	3675		3500		3350		3300	
	NB	3700		3465		3176+3573		3328	

of separate segments corresponding to film frames have been made clearer. The test word is the Finnish word *täytyypä* [tæty:pæ] 'one ought to.'

Tables 1—3 show also the results of the formant data obtained by using three spectrogram modes.

The explosions of plosives in the transitions [pæ] and [tæ] are included to the first segment of the vocoid [æ]. The duration of each segment corresponding to the frame frequency was 21,1 msec.

In the transition [pā] formants 1, 2 and 3 rise due to the low labial loci in changing over from segment *ä1* to segment *ä2*. Their frequencies are already invariable when reaching the next segment, *ä3*, but the amplitude of F1 rises from 36 to 42 db. In Fig. 3—6 the smallest distance between the labial contours grows gradually frame by frame as indicated by the following measurements: *ä1* 7 mm, *ä2* 10 mm, *ä3* 16 mm. Measuring is easy using the pointed pattern in the back plate of the apparatus; the distances of the points, which are equal in size, correspond to 4 mm of life-size of the subject's articulatory organ. The scale of the apparatus is 5-fold. Measured from the contour spectrogram this labial transition took about 40 msec. (Fig. 7).

In the transition *tä* F2 is continuously falling during the course of four successive segments as shown by the frequencies 2650—2375—2250—2125 (WB). The locus of F2 in this transition [tä] is rather high (about 2600) mostly due to the fact that the test person was very young (a 12-years old boy), but also depending on the post-dental-prealveolar position of the Finnish [t]. Examining the movements of the tip of the tongue in the corresponding frames, we find that the profile of the tongue is not completely convex and non-apical until the fifth segment, that is, around 85 msec. (4×21,1 msec. = 84,4 msec.). See Figs. 8—13. The apical distance (between the tip of the tongue and the alveolar arch) is increasing in millimeters, frame by frame, in the following way: *ä1* and *ä2* 5 mm., *ä3* 6 mm., *ä4* 7 mm., and *ä5* 14 mm.

		ä12		y1		y2	
F	Print	c/s	db	c/s	db	c/s	db
F0	NB	418		420		414	
F1	VP	720	24	425	24	425	24
	WB	575		425		425	
	NB	418 + 836		420		414	
F2	VP	2500	36	2525	36	2475	36
	WB	2525		2500		2450	
	NB	2508		2520		2484	
F3	VP	3325	36	3350	24	3325	24
	WB	3325		3350		3275	
	NB	3344		3360		3312	

In the transition of the components of the diphthong [äy] three clear stages can be distinguished. 1) F1 noticeably starts to fall (1100—950 c/s) at phase ä8—ä9. 2) F2 rises to the level of the next component [y] (2500 c/s) at phase ä10 — ä11. 3) F1 falls finally to the level demanded by [y] (400 c/s) at phase ä12—y1 (Fig. 14). In the corresponding frames we can see that the contraction of the narrowest point of the tongue channel begins at phase ä8 and ends only about phase y1 (from 21 to 13 mm). At the same time we can see the narrowest point of the pharyngeal channel extending from 11—18 mm. The length of the whole transition is thus about 5 segments, or 100 msec. The labial channel in the frames ä12, y1 and y2 has the following measurements: 24 mm in the last [ä] frame, 20 mm in the first and 8 mm in the second [y] frame (Figs. 15—21).

Without explaining the corresponding acoustic phenomena, I shall show some diagrams made from a film where the rate of 80 frames per sec. was used (Fig. 23). We shall compare the last frames of the preceding sounds with the first ones of the later. In Fig. 24 the tip of the tongue has not yet quite reached the alveolar arch as the vowel [e] is just changing into the Finnish medioalveolar [s] sound in the test word *Esko* (a Finnish male name). The narrow opening of the front teeth and the position of the mediodorsum of the tongue are already the same as in the next frame 12,5 msec. later (Fig. 25). The transition of [es] during two successive frames, that is 25 msec., is demonstrated in Fig. 26 where the two phases of movement are photographed simultaneously on one and the same diagram. The CC-transition [sk] of the same test word is shown in Fig. 27 as above in two successive phases of transition. One can see the tongue preparing itself to the mediopalatal occlusion of [k] in the frame of the last phase of [s]; the slight upward movement (5 mm) of the back of the tongue and at the same time the quick release movement (4 mm) of the tip very skilfully produce the transition from [s] to [k]. The jaw has hardly had time to change its position at all during this dual picture.

The following diagram (u)r + (u)r2 (Fig. 28) concerns this same [r] contoid, in which the position (u)r2 corresponds to the wider phase in the tremulation of the front part of the tongue. The narrower phase of the tremulation has been diagrammed from the preceding frame. The duration of one vibration of the tongue is about 25 msec. and the tremulation has a frequency of about 40/sec. in this contoid. The narrower phase resembles an apical medioalveolar occlusion as in the Finnish plosive [d], but in this phase the air flow is going through a very narrow groove.

In the near future we will investigate the preceptual level of those transitions of frame diagrams and spectrograms. For this purpose we will use our transistorized photocell segmentator listening to the desired cut sequences of the optical sound of the X-ray motion pictures.

LITERATURE

- Delattre, Pierre (1965). De la hiérarchie des indices acoustiques pour la perception des consonnes. *Proceed. of the Fifth International Congress of Phonetic Sciences* (Münster 1964), p. 244—251.
- Fant, Gunnar (1965). Formants and Cavities. *Proceed. of the Fifth International Congress of Phonetic Sciences* (Münster 1964), p. 120—141.
- Ondráčková, Jana (1965). On the Roentgenography of the Speech Organs. *Proceed. of the Fifth International Congress of Phonetic Sciences* (Münster 1964), p. 446—449.
- Sovijärvi, Antti (1962). Röntgenkinematographisch-akustische Untersuchungen über die Artikulation der Diphthonge. *Proceed. of the Fourth International Congress of Phonetic Sciences* (Helsinki 1961), p. 111—128. Ibid. 1963: Suomen kielen äännekuvasto. Jyväskylä.
- Subtelný, Joanne D. & Subtelný, J. Daniel (1962). Röntgenographic Techniques and Phonetic Research. *Proceed. of the Fourth International Congress of Phonetic Sciences* (Helsinki 1961), p. 129—146.

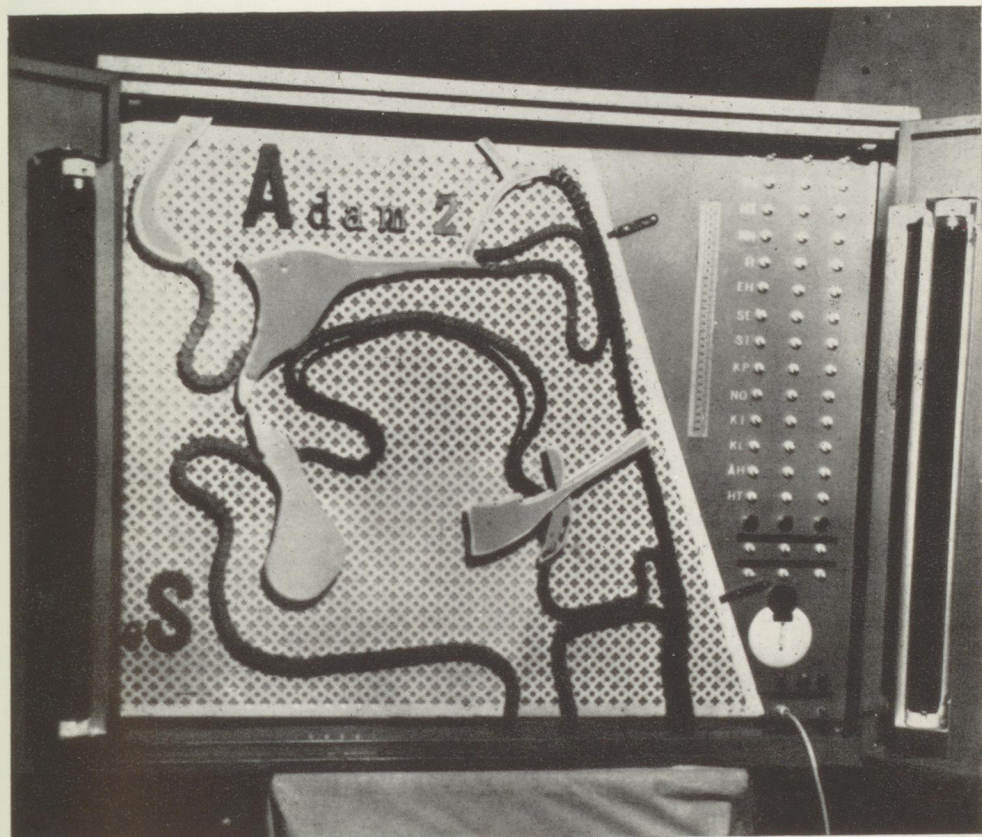


Fig. 1.

Sovijärvi: On transition in the light of X-ray films

VOICEPRINT LABORATORIES - P. O. BOX 835 - SOMERVILLE, NEW JERSEY

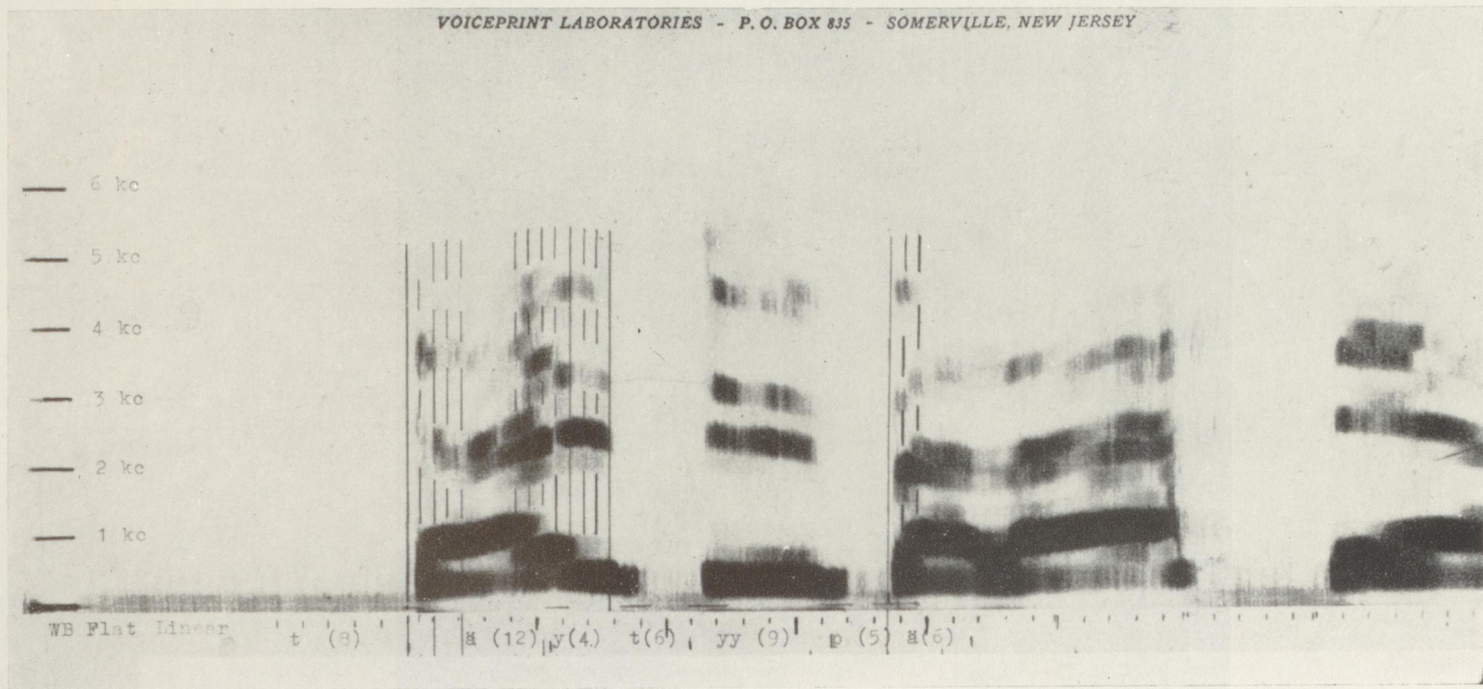


Fig. 2.



Fig. 3.



Fig. 4.

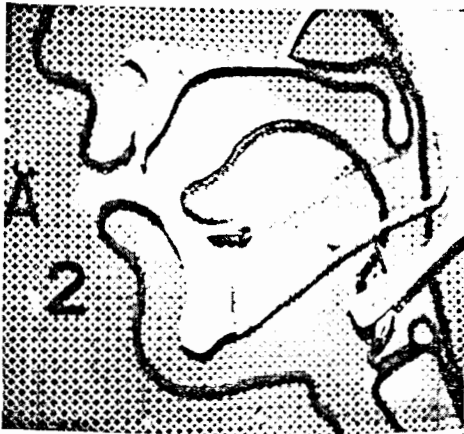


Fig. 5.



Fig. 6.

Sovijärvi: On transition in the light of X-ray films

VOICEPRINT LABORATORIES - P. O. BOX 835 - SOMERVILLE, NEW JERSEY

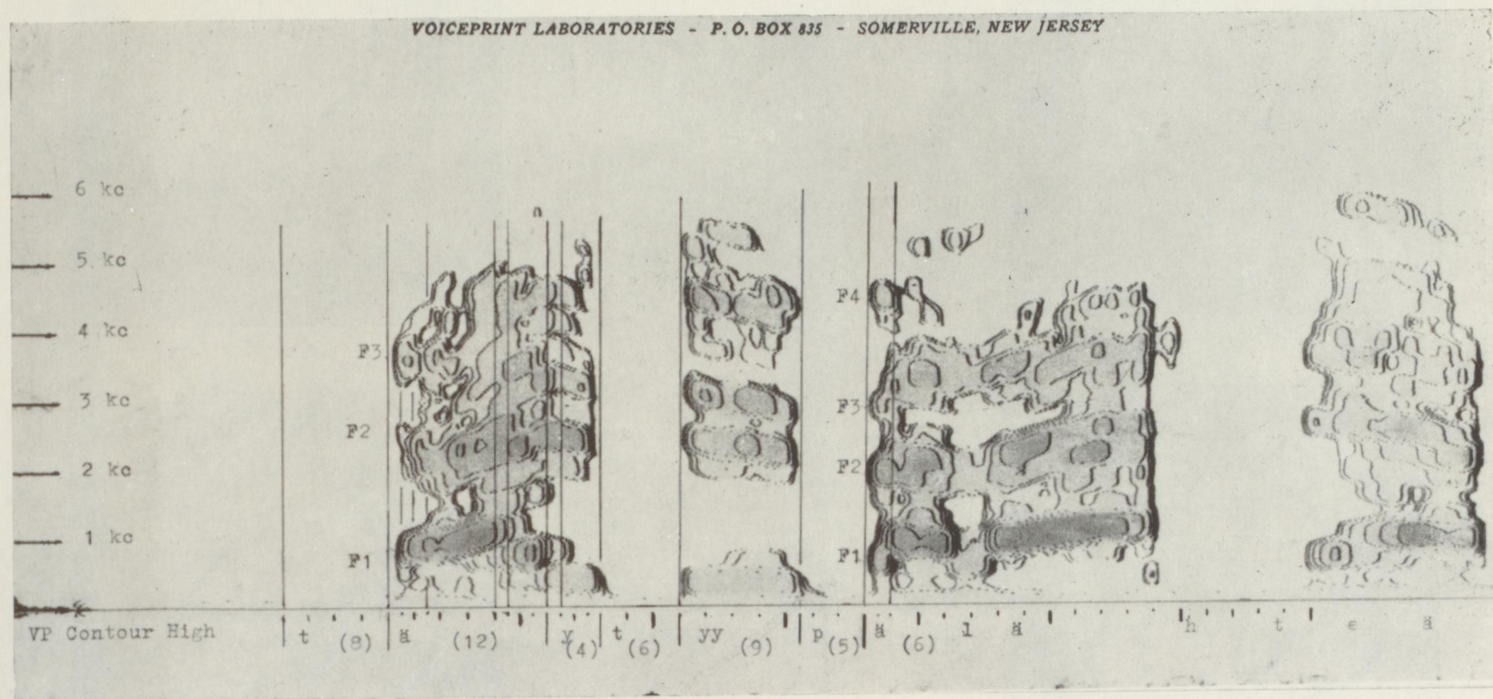


Fig. 7.

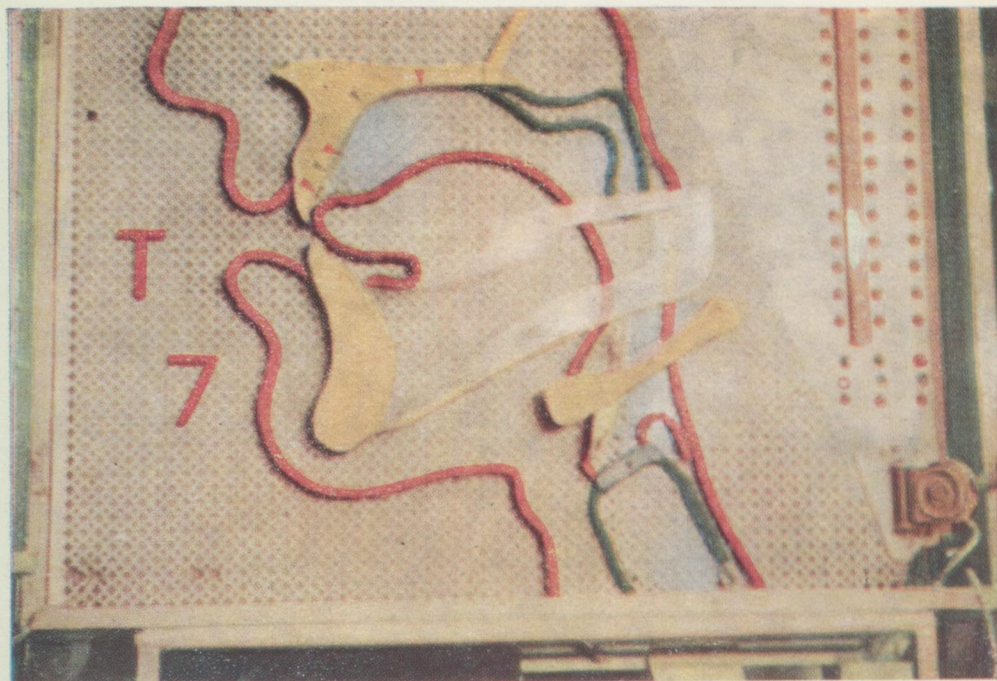


Fig. 8

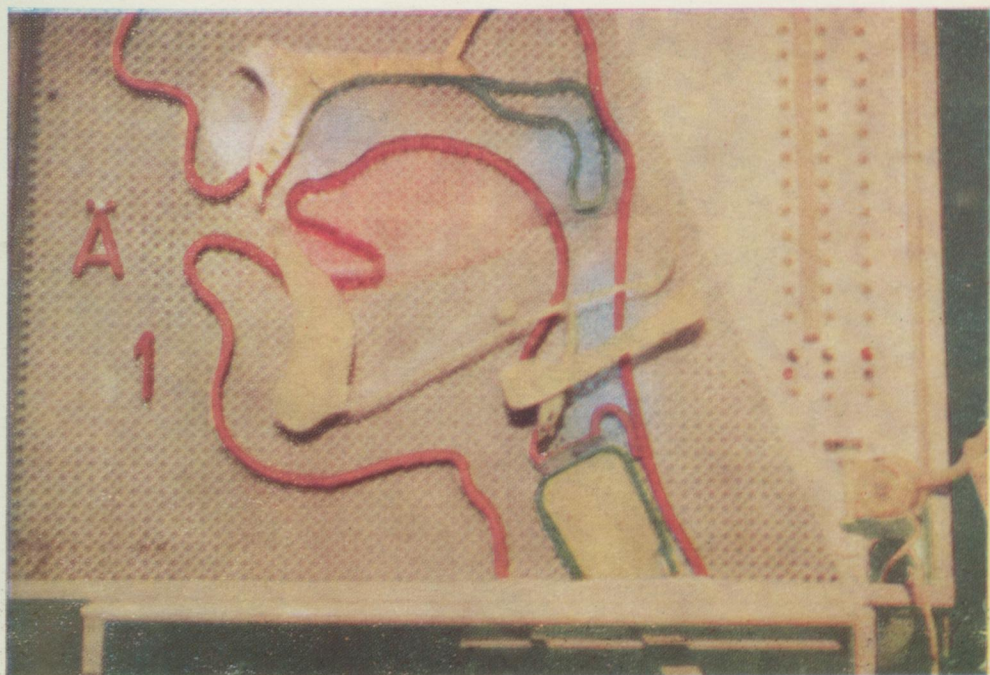


Fig. 9

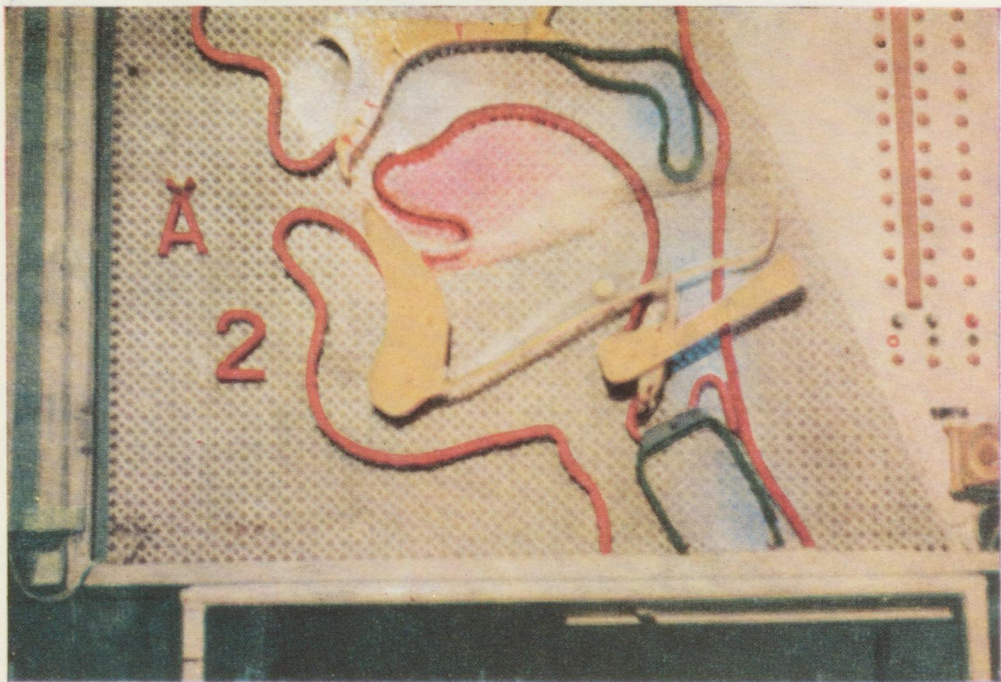


Fig. 10

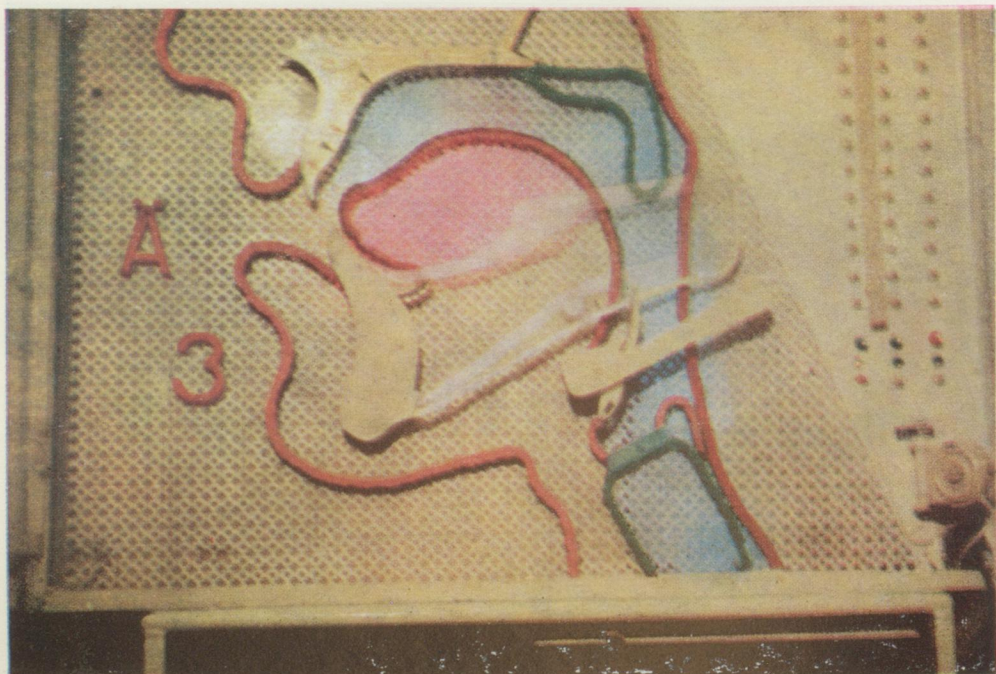


Fig. 11

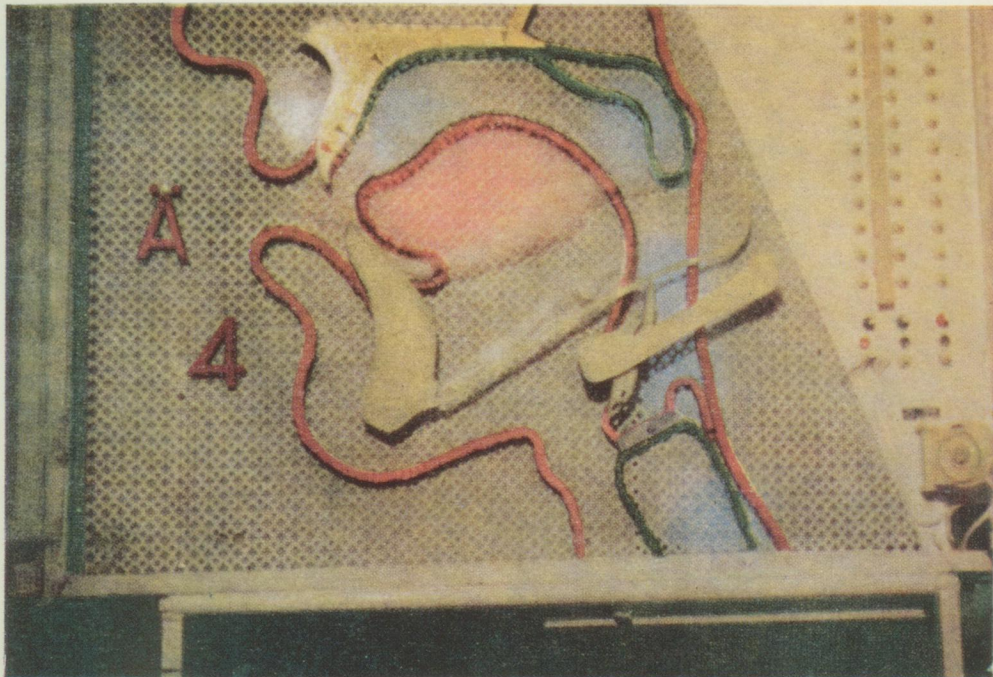


Fig. 12

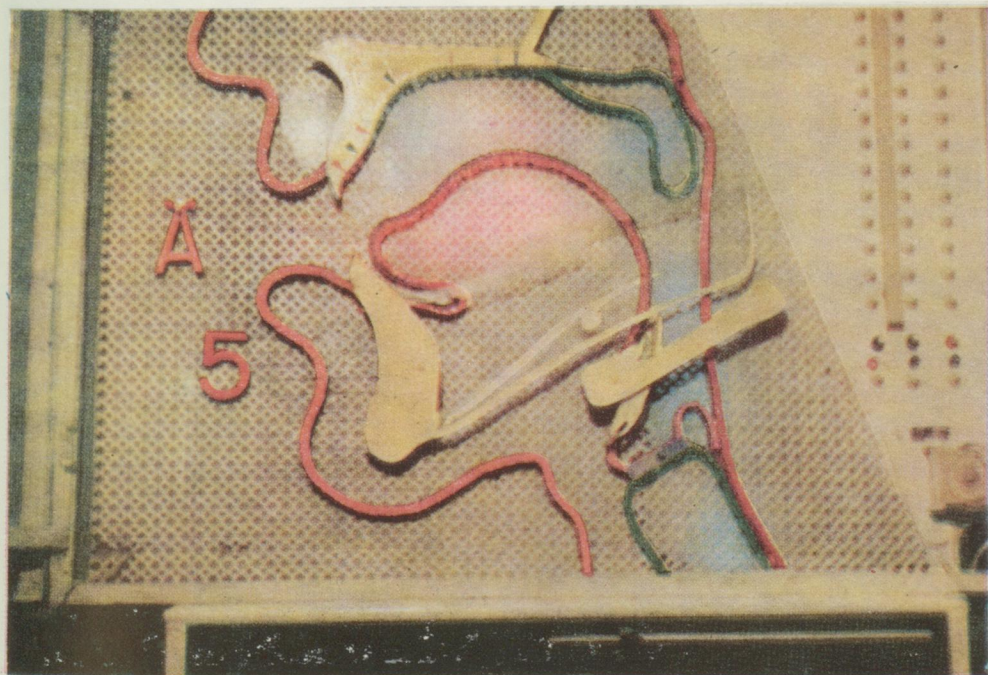


Fig. 13

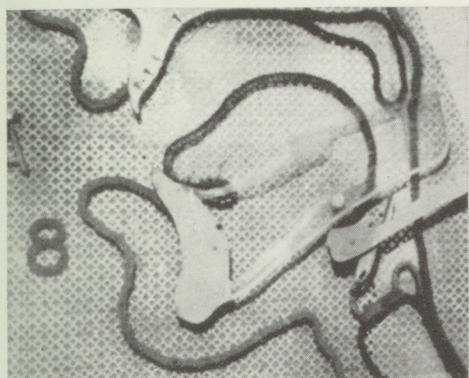


Fig. 14.

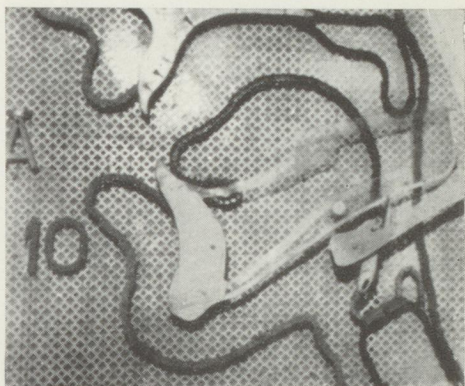


Fig. 15.

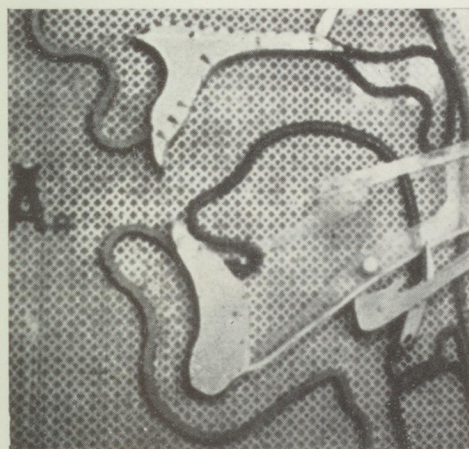


Fig. 16.

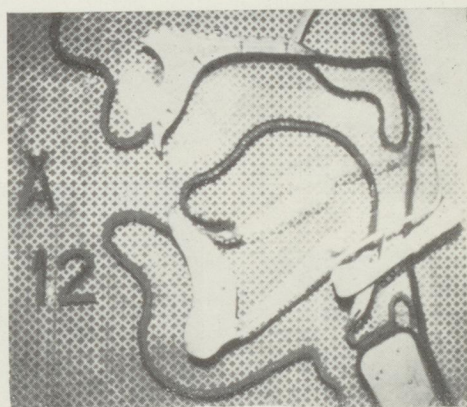


Fig. 17.

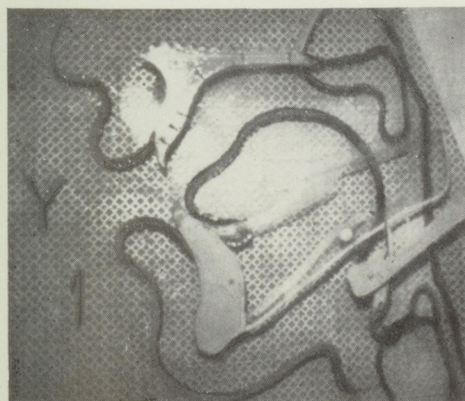


Fig. 18.

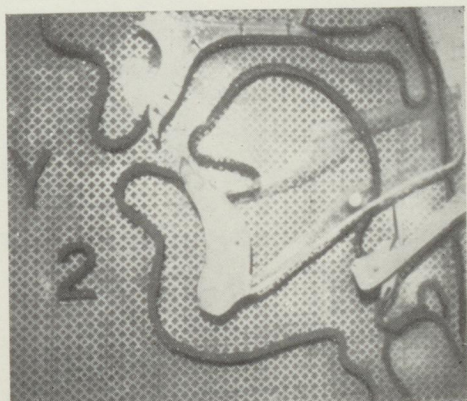


Fig. 19.

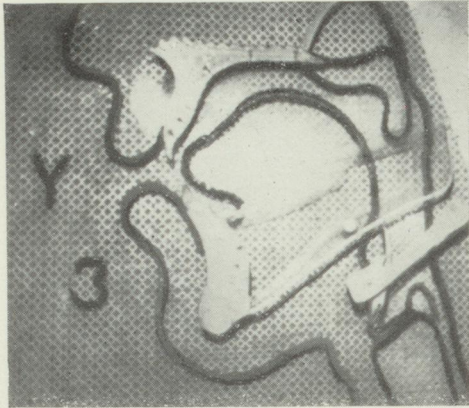


Fig. 20.

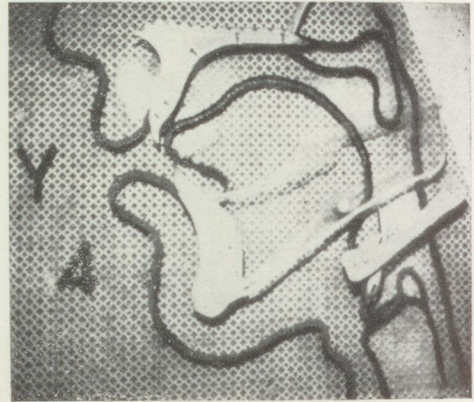


Fig. 21.

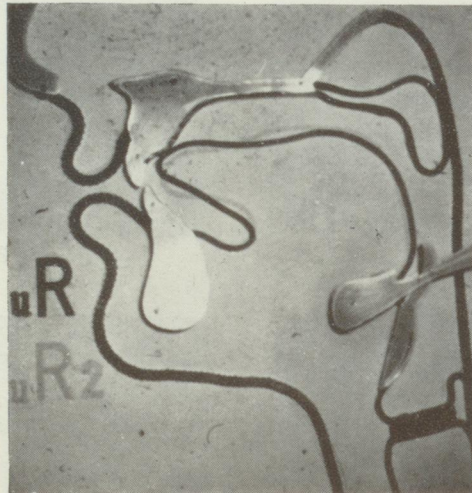


Fig. 28.

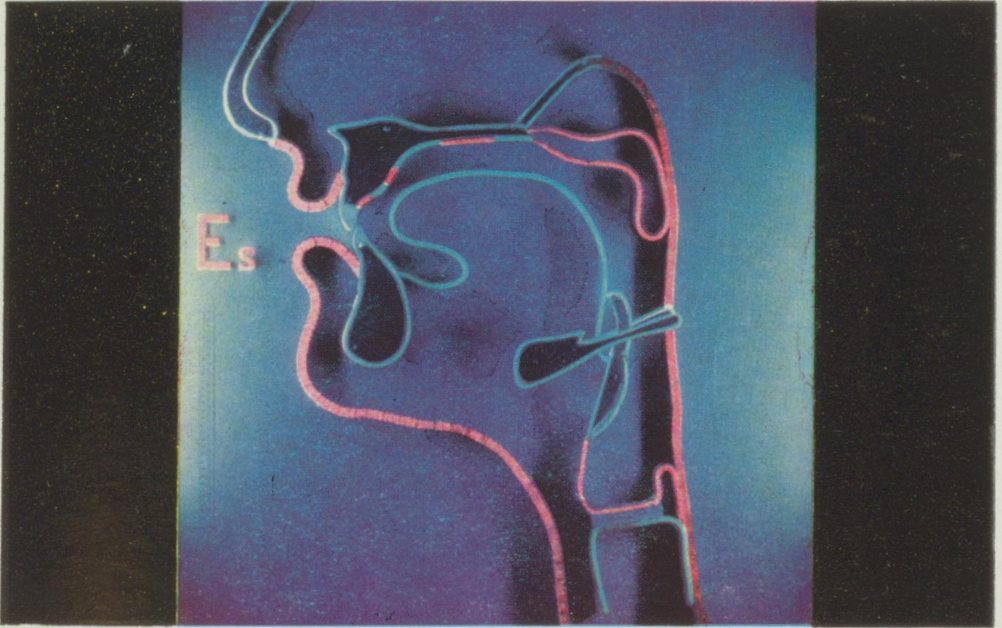


Fig. 22

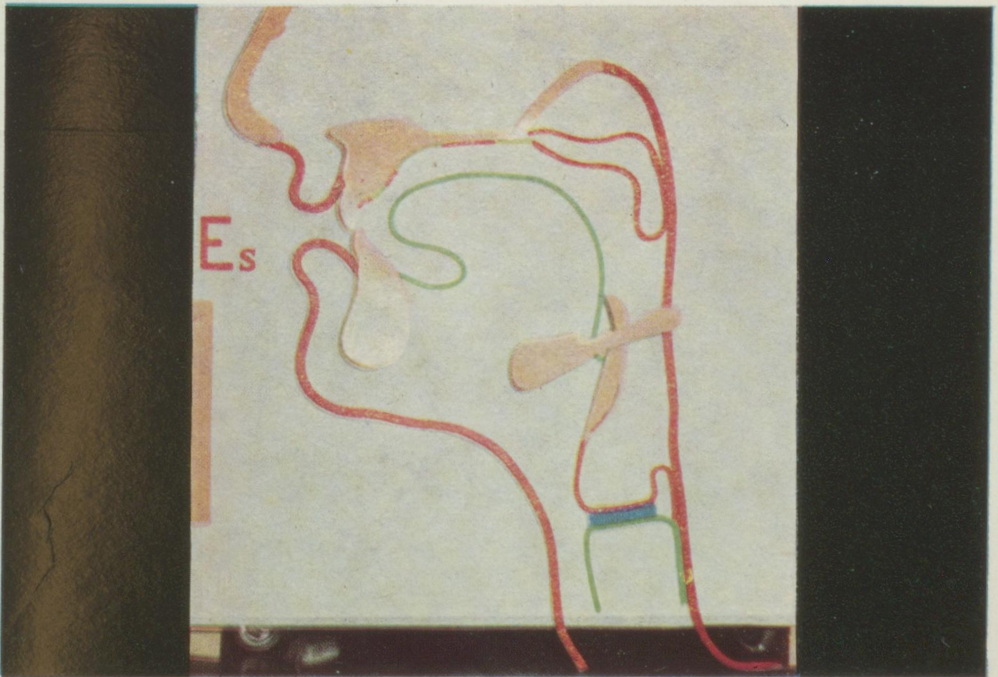


Fig. 23

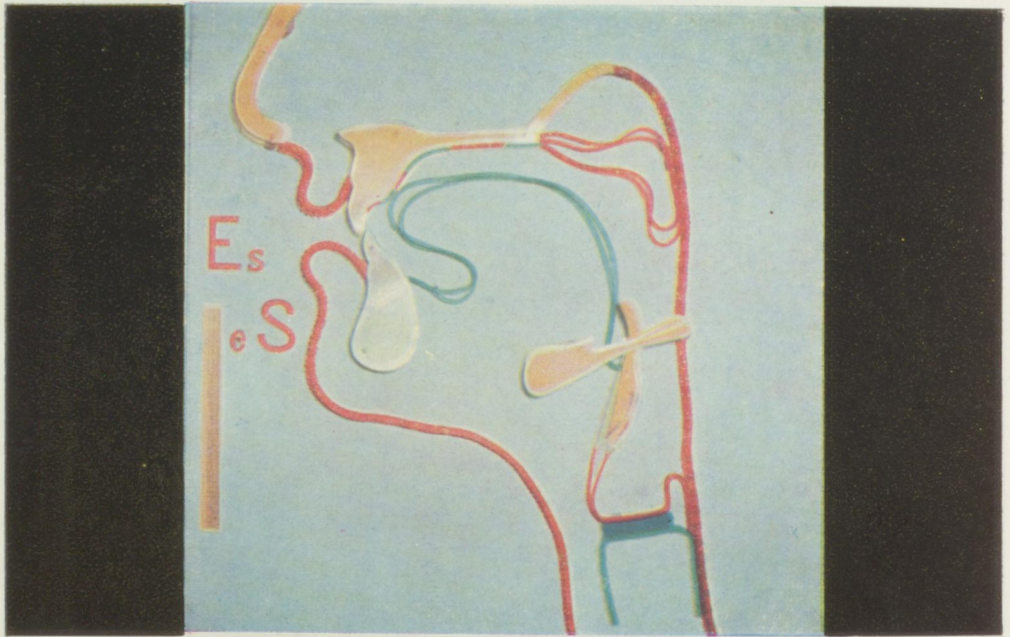


Fig. 24

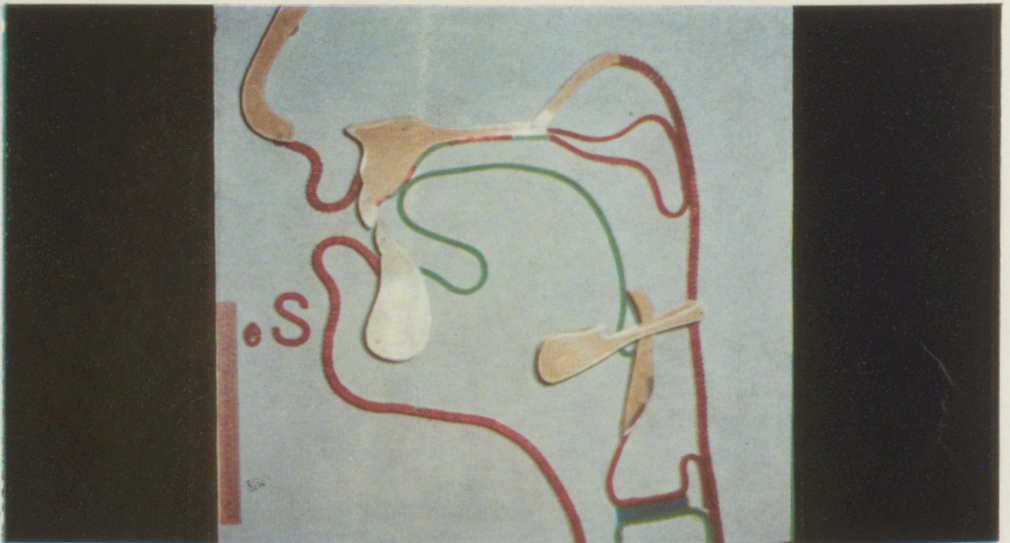


Fig. 25

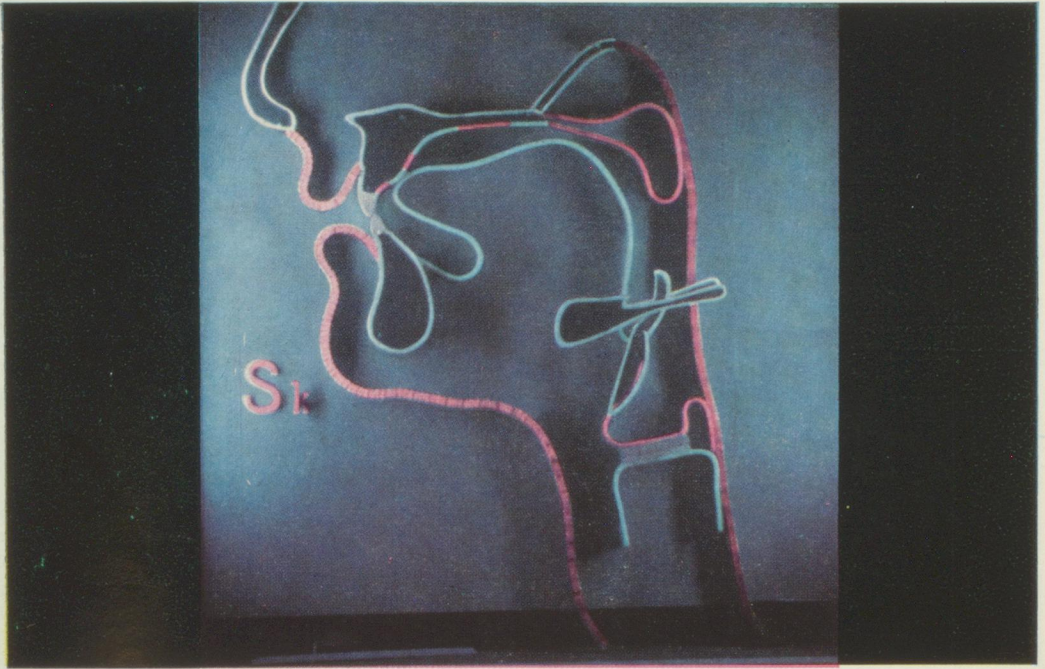


Fig. 26

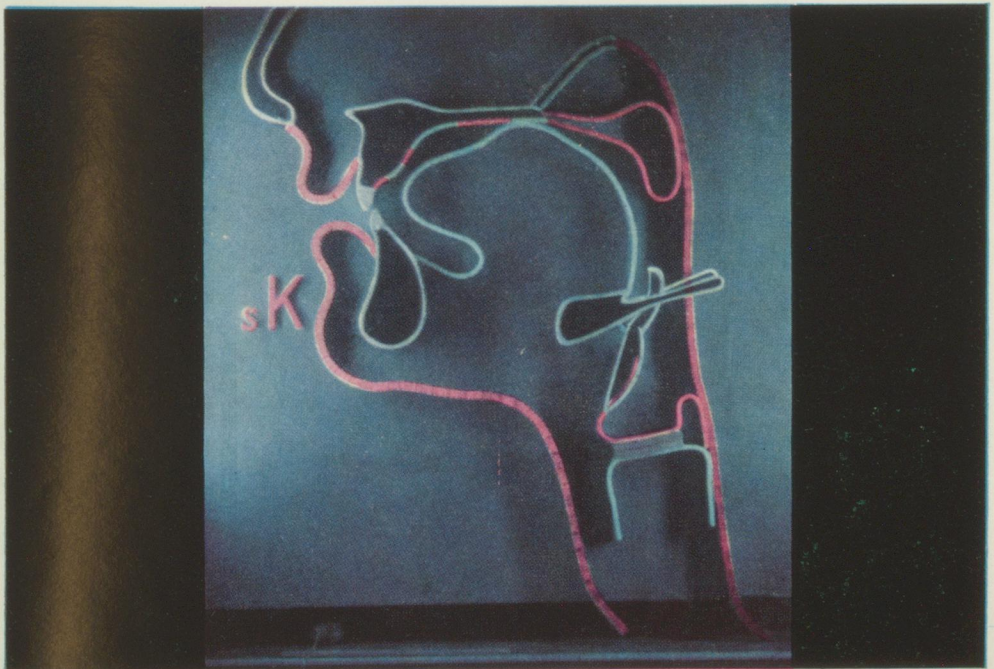


Fig. 27

EXAMPLES OF X-RAY FILMS FROM FINNISH (DEMONSTRATION)

ANTTI SOVIJÄRVI

- a) At first we shall see the X-ray sound film (from the year 1961) which has been taken with 48 pictures/sec. It has a synchronized stretched sound made by means of the "Temporegler" in the Finnish Broadcasting Co.
- b) The first word of this film *täytyypä* [täyty:pæ] has been diagrammed by means of our demonstrating apparatus ADAM. There are different variations of this trick film. The refilming is made picture by picture in 1963 and 1966.*
- c) We used the speed of 60 pictures/sec. and put a wedgeshaped metal plate behind the area of the front part of the subjects' vocal tract. Dynamic variants of stressed syllables have been investigated in this film (from the year 1965) (Fig. 1).
- d) 80 pictures/sec. has the next example of our series made in 1966. We fitted *several* smaller metal plates behind the labial and laryngeal area. Contrasting materials have not been further used (Fig. 2).
- e) In this year we made the invention of a plastically *modelled* alloy filter which is fitted on the speakers shoulder (Fig. 3). The thickest parts of this filter correspond to the labial and laryngeal areas, but the thinnest parts to the velic and sublingual areas. We have filmed apical [r] sounds and plosives in Finnish test-words spoken by two female assistants. The speed has been again 80 pictures/sec. The slowing to normal motion is therefore 3,3-fold. This film is without sound. (Fig. 4)

The data for filming our newest film were following:

1. The shooting time per frame 4 msec.,
2. the amplification: 80 kV; 125 mA,
3. the grade of "Schwärzung" 0,75, and
4. GPan 30, 20 din.

The newest film has been made as a result of the team work of Ass. Prof. E. Tähti (Helsinki), Prof. F. E. Stieve (Munich), Mr. A. Steiner (Berlin), Mr. N. Jaronen (Helsinki), and myself at the Central Hospital of Helsinki University Aug. 24, 1967. The construction of the modelled metal filter of X-rays is made by our technician Mr. Jaronen and myself.

* Cf. fig. 15—21 in A. Sovijärvi, On transition in the light of X-ray films (this volume, p. 856 f.).

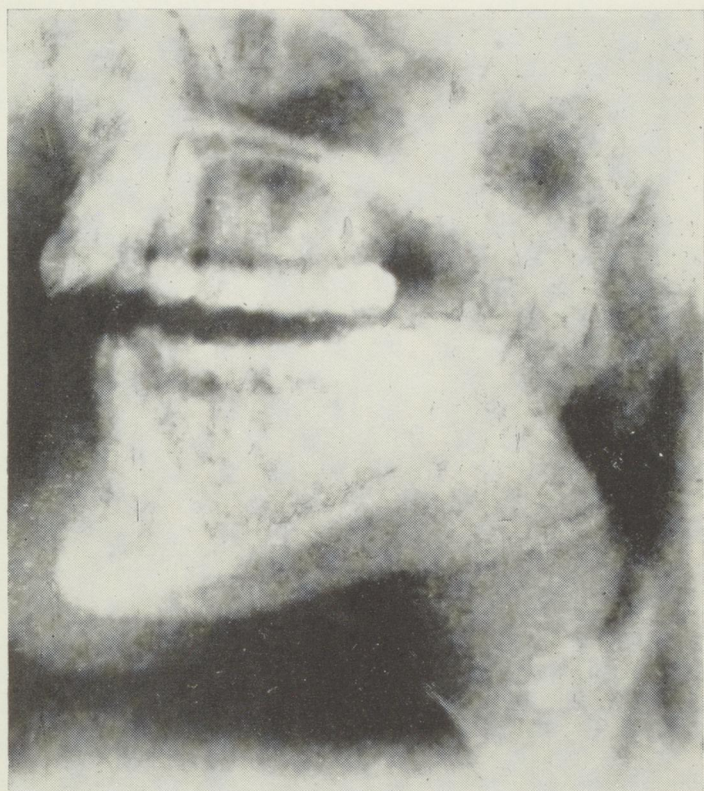


Fig. 1.



Fig. 2.

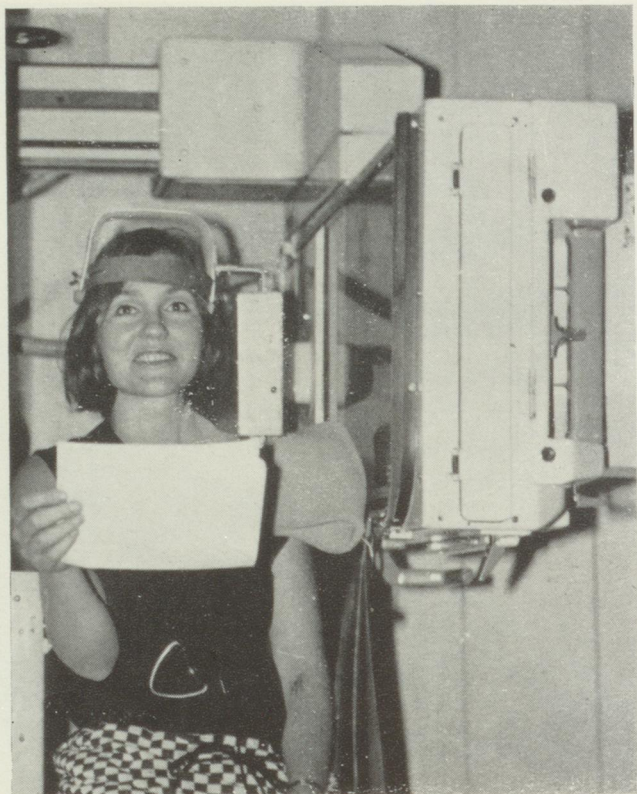


Fig. 3.

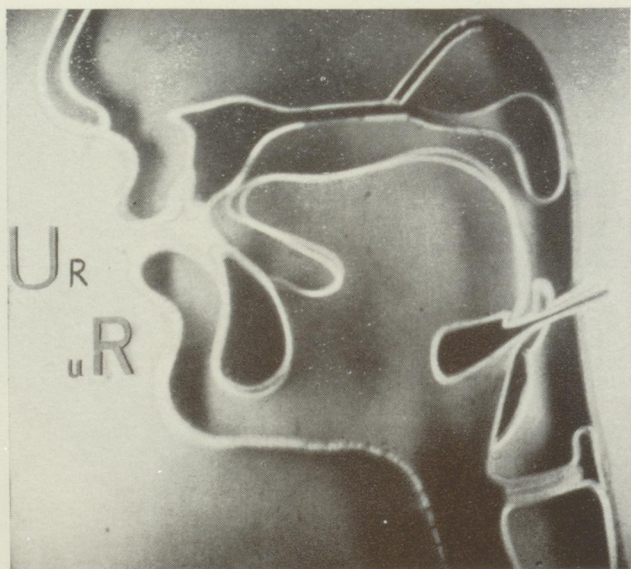


Fig. 4. The physiological transition of *ur* in the word *purkki*.

Ur = the last frame of the vowel *u*; uR = the first frame of the apical tremulant *r*.

The duration of each frame 12,5 millisecc. = 1,25 sentisecc. The speed has been 80 pictures/sec.

DIE BEDEUTUNG DER AUDITIVEN KONTROLLFUNKTION FÜR DIE STEUERUNG VON SPRECHSITUATIONEN ALS TYPISCHEN INFORMATIONSEINHEITEN

H. STELZIG*

Gesprochene Sprache als dominierendes Kommunikationsmittel verfügt über ein Informationsinventar von verschiedenartig bedingten, differenzierten Sprechsituationen. Es sind Sinneinheiten mit spezifischem Informationswert; sie resultieren propriozeptiv als Aktion oder rezeptiv als Re-Aktion aus dem jeweiligen Milieu, in dem der Sprecher lebt und zu dessen Bestand oder Veränderung er durch einen zielgerichteten Sprechakt beiträgt. Als objektive Widerspiegelungen des Verhaltens zur Umwelt im Ausdrucksfeld der mündlichen Kommunikation sind sie mit bestimmten Milieufaktoren koordinierbar und anpassungsfähig (z. B. Raumverhältnisse, Hörerkontakt). Art und Formung der Sprechsituationen erfolgt differenzierend nach den Mitteilungsbedürfnissen, die einem hochentwickelten sozialen Kontakt dienen. Sie lassen sich bestimmten neurophysiologischen Kategorien zuordnen, die als dominierende Steuerungszentralen fungieren: die trophotrope und die ergotrope Funktionsrichtung dem Vegetativum, die ideotrope dem Kortex. Sprechsituationen mit vegetativer Dominanz erfordern bei präzisierter Information im Gestaltungsablauf größere kortikale Steuerung und Kontrolle; bei Affektausdruck reduziert sich dementsprechend der emotionale Anteil oder er wird poetisch stilisiert. Mit der komplexen Steuerung von Sprechsituationen wird bereits das gesamte Energiepotential des Sprechvollzugs im Bereich der drei physiologischen Funktionskreise (z. B. Atem- und Stimmvolumen, Artikulationsspannung) eingesteuert; die Sprechspannung ist der spezifischen Sprechsituation proportional. Es ist ein geschlossener Funktionskreis von Anregung (Reizauslösung) — Ausführung (Verarbeitung über den motorischen Prozeß) — Auswirkung (Hör- und Erkenntnisprozeß). Innerhalb der genannten Funktionsrichtungen ergibt eine Ermittlung nach distinktiven Aspekten ein System von Grundmodellen, die Sprechsituationen als typische Ausdruckseinheiten ausweisen. So gehören z. B. zur trophotropen Biofunktion folgende Grundmodelle: Sympathie, Beruhigung, Trost, Sehnsucht, Bitte, u. a.; zur ergotropen Richtung zählen: Stolz, Zorn, Trotz, Furcht, u. a.; als ideotrop gilt allgemein affektfreier Ausdruck. Die Differenzierung im Sprechprozeß läßt die distinktiven Merkmale der Grundmodelle deutlich erkennen; dadurch gewinnen sie typischen Signalwert

* Universität Greifswald/DDR, Abt. Sprechwissenschaft.

und dienen der Motivation im Sprechausdruck. Dem Kollektivcharakter der Sprache entsprechen Kombinationsmöglichkeiten von unterschiedlichen Sprechsituationen, die im wesentlichen durch Assoziierung (z. B. Trauer : Trost), Vergleich und Steigerung (Neid : Haß z. B.) oder Frage und Antwort (z. B. didaktisches Gespräch) näher bestimmbar sind.

Der Sollwert eines Grundmodells ist konkretisierbar; für die Realisierung sind semantische, linguistische und phonetische Kreisprozesse von grundlegender Bedeutung. Die Speicherpotenz für Grundmodelle der Sprechsituationen ist vom Perzeptionsorgan des Sprechers abhängig. Jedes Modell wird innerhalb einer großen situationsgemäßen und individuell bedingten Variationsbreite realisiert, es bleibt aber in seiner Eigenart erkennbar und ist somit lehrbar und erlernbar. Da der Istwert der jeweiligen Grundmodelle psychophysisch steuerbar ist, läßt sich die Konstanz und kommunikative Wirksamkeit der Sprechsituationen gewährleisten. Die Einstellung auf die erforderliche Stellgröße eines Grundmodells wird durch Reizfolgen ausgelöst. Zwischen Sollwert und Istwert bestehen Ausgleichsmöglichkeiten über die Rückkopplungsfähigkeit der Regelstrecken, die im allgemeinen als Kontrollfunktion zu wenig ausgebildet und genutzt wird, um dem Sprechausdruck einen optimalen Wirkungsradius zu sichern. Über die kybernetischen Funktionen bei der Stimmgebung und beim Sprechen ist nach den Untersuchungen von Winckel eine laufende Kontrolle und Steuerung im mündlichen Ausdrucksbereich möglich. Im normalen Sprechprozeß bilden die sprechmotorischen und sprechakustischen Funktionen eine einheitliche Assoziationskette, wobei die auditiven Parameter auf Mittelwerte mit normaler Schwankungsbreite orientieren.

Die Realisierung emotional bestimmter Grundmodelle unterliegt einer größeren Schwankungsbreite als konventionelle Muster im normalen Ausdrucksfeld. Aber auch im vegetativ gesteuerten Sprechprozeß ist die Schwankungsbreite begrenzt durch den Kollektivcharakter der Sprache. Folgende Störgrößen ergeben Diskrepanzen zwischen Sollwert und Istwert:

1. ungenügende Reizauslösung, die physisch wie auch psychisch bedingt sein kann;
2. mangelhafte Speicherung von verfügbaren Grundmodellen für Sprechsituationen, vor allem im Bereich der kortikalen Kategorie und somit bleibt bereits die Bereichseinstellung auf den Sollwert unklar. Bei gesprochener Dichtung als Nachvollzug gilt es für die vegetativen Kategorien;
3. ungenügende Kontrolle über die Regelstrecken, somit erreicht der Istwert nicht den Sollwert und die kommunikative Wirkung wird reduziert.

Unterschwellige Ausdrucksgestaltung, die vor allem durch mangelnden Wortschatz oder ungenügende rhetorische Erfahrung bedingt sein kann, realisiert das Ziel nur teilweise. Überschwellige Gestaltung wirkt unglaubwürdig, pathetisierend. Bei vegetativ bedingten Grundmodellen führt eine kortikale Übersteuerung zu Rationalisierung und Nivellierung des Gefühlsanteils. Zu starke vegetative Steuerung ergibt Individualisierung und Subjektivierung.

Diese Störgrößen lassen sich durch die Ausbildung der auditiven Kontrollfunktion und die Orientierung auf Grundmodelle, die eine Lernmotivation erleichtern, weitgehend ausschalten. Sprechsituationen als übergeordnete Einheiten bestimmen auch die phonetischen Parameter.

LITERATUR

- Trojan, F.: *Der Ausdruck der Sprechstimme*, Wien/Düsseldorf 1952.
Winckel, F.: Kybernetische Funktionen bei der Stimmgebung und beim Sprechen. *Phonetica* 9 (1963), 108—126.
Stelzig, H.: Grundlagen der gesprochenen Dichtung. Habil. Schrift Greifswald 1965 (hier alle weiteren lit. Hinweise).

DISCUSSION

Geissner:

Wenn der Referent sagte „Sprechspannung ist der Sprechsituation proportional“, dann setzt das einen zu einfachen Reiz-Reaktionsmechanismus voraus, bei dem sie Situationsmodelle als invariante Einheiten betrachtet werden. Da diese Annahme vielen sozialpsychologischen und soziologischen Ergebnissen auf diesem Gebiet zu widersprechen scheint, ließe sich nur sagen: „Sprechspannung soll der Sprechsituation proportional sein“. Dies ist aber ein pädagogisches Postulat, weniger ein analytisches.

A THEORY OF THE ORIGIN OF AKAN VOWEL HARMONY

J. M. STEWART*

Akan (Twi-Fante), the dominant indigenous language of Ghana, is one of the many African languages which have vowel harmony based on the so-called 'lax/tense' distinction: as a general rule words consisting of a one-morpheme stem with or without prefixes have either 'lax' or 'tense' vowels throughout, e.g. *nŏm* "drink", *nŏm* "suck", *wbenŏm* "you will drink", *wbenŏm* "you will suck". In recent years Ladefoged (1964), myself (1966) and Pike (1967) have expressed the view that the so-called 'lax/tense' distinction, in these African languages at least, is not so much a matter of laxness *vs.* tenseness as of unadvanced *vs.* advanced position of the tongue root. In this paper I consider an Akan problem which raises a further general question, namely that of the relationship of the so-called 'lax/tense' distinction to the voiceless/voiced distinction.

There is reason to suspect that Akan vowel harmony has its origin in the voiceless/voiced distinction. As I have shown elsewhere (1967), the consonant system has the distinctions *p/t/k/kw*, voiceless/voiced, and occlusive/continuant, and the vowel system the distinctions *i/u/e/o/a*, oral/nasal, and root-unadvanced/root-advanced ('lax/tense'), but there is no voiceless/voiced or occlusive/continuant distinction in final consonants and this, taken together with other facts, suggests a theory on the following lines: (i) at one time final consonants had the full range of consonantal distinctions and there was no root-unadvanced/root-advanced ('lax/tense') or oral/nasal distinction anywhere; (ii) *either* final voiced consonants *or* final voiceless consonants became spontaneously root-advanced ('tense'), while final occlusive consonants became spontaneously nasal; (iii) vowels became root-advanced ('tense') before root-advanced ('tense') consonants and nasal before nasal consonants; and (iv) a number of changes eroded the voiceless/voiced and occlusive/continuant distinctions in final position (although the nasal occlusive *n* and the oral continuant *r* survived as conditioned variants of the final alveolar consonant after nasal and oral vowels respectively). This suggests in turn that the vowel harmony came about as a result of the root-advancing ('tenseness') which emerged spontaneously in certain final consonants spreading not only to the preceding vowel but to the entire preceding part of the word.

*Institute of African Studies, University of Ghana.

The problem is whether the root-advancing ('tenseness') is more likely to have arisen in the voiceless or in the voiced consonants. Jakobsonian phonologists would presumably consider it more likely to have arisen in the voiceless consonants as they identify 'tenseness' in vowels with the 'fortis' feature in consonants and the 'fortis' and voiceless features commonly occur together. Their identification of the 'tense' and 'fortis' features, however, is derived from their view of 'tenseness' as actual tenseness, whereas in the light of recent work on African languages, as has been seen, it seems to be not so much tenseness as advancing of the tongue root. Now tongue root advancing is more likely to have arisen in the voiced than in the voiceless consonants, since in voiced consonants the throat cavity is sometimes enlarged to allow a greater flow of air for the vibration of the vocal cords (see Hudgins and Stetson 1935) and tongue root advancing is one of the ways of enlarging it.

The plausibility of this refinement of the theory can be demonstrated by instances in present-day languages of root-advanced ('tense') vowels occurring to the exclusion of root-unadvanced ('lax') vowels in the vicinity of voiced consonants. There is in fact one instance in Akan itself; in southern dialects the root-advanced ('tense') vowels *i*, *u*, *e*, *o* generally occur to the exclusion of their root-unadvanced ('lax') counterparts *ɪ*, *ʊ*, *ɛ*, *ɔ* after 'compact' voiced stops, e.g. *jɪmɪ* "be stupid", *gu* "cast down", *je* "accept", *jɛɛŋ* "think", *gorɔ̃* "play", *jɔ̃o* "become cool".

There is another instance in my own idiolect of the Aberdeenshire subdialect of the Scottish dialect of English. Word-final unstressed '-ie', which frequently constitutes a diminutive suffix, has at least the two variants *ɛ*, *i*. It is invariably *i* after a stressed syllable with *i*, *u*, *æi*, *ɔu*, e.g. 'jellie' *jɪli* "jelly", 'hoosie' *husi* "house", 'wifie' *wɪɪfi* "woman", 'dowie' *dɔ̃ui* "sad"; cf. 'lassie' *lʌsɛ* "girl", 'berrie' *berɛ* "berry", 'bonnie' *bɔ̃nɛ* "beautiful". It is also invariably *i* after a voiced stop or fricative, e.g. 'laddie' *ladi*, 'bedie' *bɛdi*, 'froggie' *fɾɔgi*; compare the last example with 'frockie' *fɾɔkɛ*.

Yet another instance is provided by standard English pairs such as 'breath' *brɛθ*, 'breathe' *brið*. It is true that here the 'tenseness' of the vowel and the voicing of the final consonant are not historically related, but the similarity of "laxness" and 'tenseness' to voicelessness and voicing respectively may well have been a factor favouring the survival of the vowel alternation.

REFERENCES

- Hudgins, C. V. and R. H. Stetson (1935): 'Voicing of consonants by depression of the larynx.' *Arch. Néerland. Phon. Expér.*, 11, 1—28.
- Ladefoged, P. (1964): *A Phonetic Study of West African Languages*, Cambridge University Press.
- Pike, K. L. (1967) 'Tongue root position in practical phonetics.' *Phonetica* (forthcoming).
- Stewart J. M. (1966): 'Tongue root position in Akan vowel harmony.' *Phonetica* (forthcoming). (First presented at the Sixth West African Languages Congress, Yaoundé, 1966.)
- Stewart, J. M. (1967): 'A deep phonology of the Akan monosyllabic stem.' *Journal of West African Languages* (forthcoming). (First presented at the Seventh West African Languages Congress, Lagos, 1967.)

DISCUSSION

Brend:

I trust it was clear from my paper that the spectrographic measurements I made did seem to confirm your postulation of tongue-front position for the one set of vowels. The expected acoustic result, i.e. the lowering of the first formant, did regularly occur.

Carnochan:

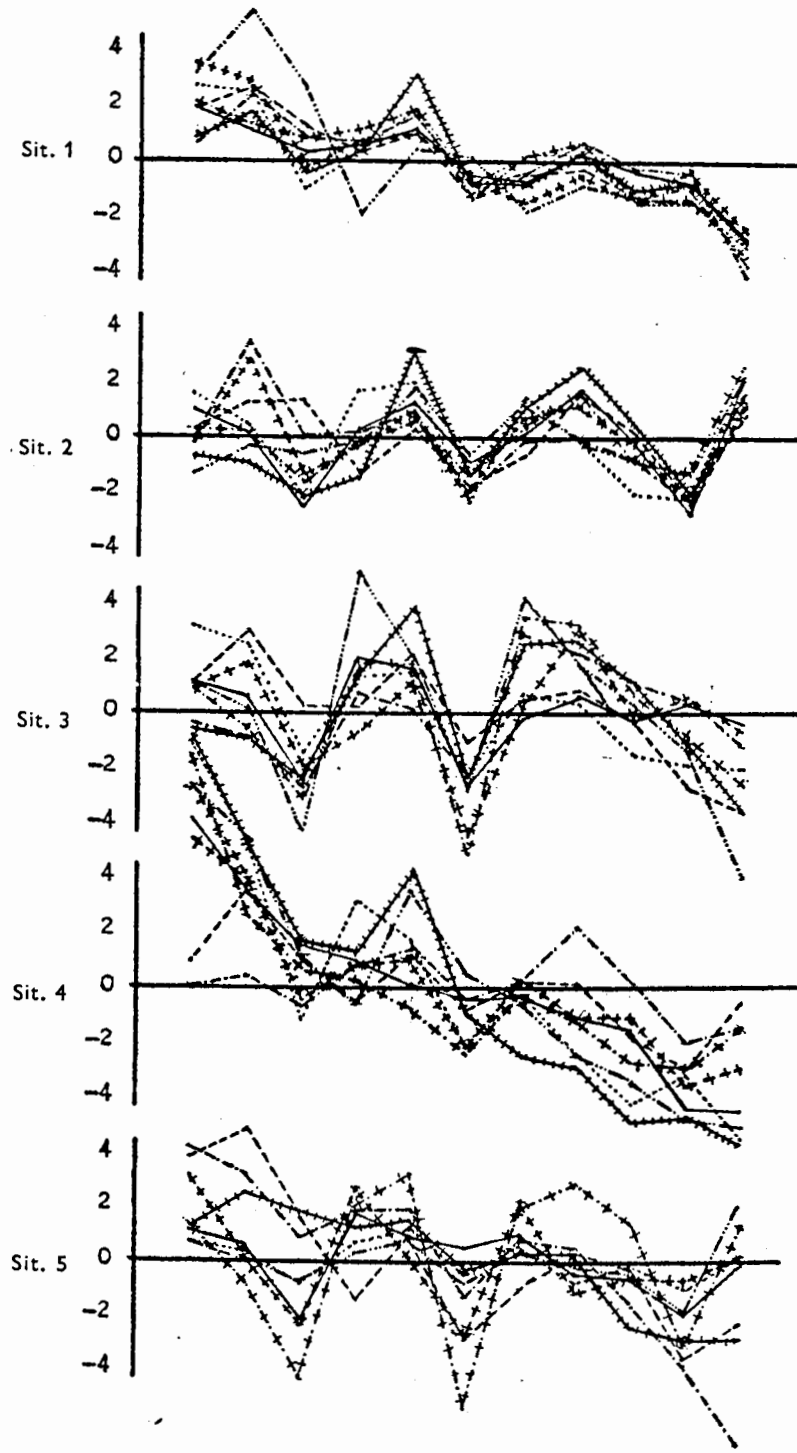
It is interesting to suggest that different positions of the root of the tongue are associated with the tense-lax distinction sometimes described for vowel-harmony. I welcome this for Igbo, as it would provide a physiological similarity of articulation of the vowels in successive syllables, for which the tongue positions are otherwise quite different, thus offering additional phonetic exponents for the R/L prosodic elements of structure already established.

ZUR BEWERTUNG DES TONHÖHENVERLAUFS IM AUSSAGESATZ

E. STOCK

Fragestellung. Es steht außer Zweifel, daß ein Melodem vor allem durch die Abstufung der sog. Führtöne (v. Essen) charakterisiert wird. Offen bleibt, bis zu welchem Grade die Grundtonbewegung in nichtiktischen Silben berücksichtigt werden muß. Diese Frage ergab sich für uns aus Messungen, die wir an Tonhöhen-
diagrammen (Stock) von über 500 Aussagesätzen durchgeführt haben. Dabei wurde die Grundtonbewegung im Silbenkern ausgemessen und gemittelt; Gleittöne (Maack, Elekfi u. a.) blieben unbeachtet. Die sich aus diesen Meßwerten ergebenden Kurven lassen erkennen, daß nichtiktische Silben melodisch häufig nicht an die Iktussilben pro- oder enklitisch angelehnt werden (s. Abb. 1: Sit. 1: sachlich, 2: verbindlich, 3: nachdrücklich, 4: verächtlich, 5: zornig). In einer Reihe von Fällen sind Unterschiede zwischen den Mittelwerten der Silbenmelodie (kurz *Differenzen* genannt) so auffällig, daß eine Kontrastverschärfung angenommen werden könnte. Ob solche Differenzen in der ektosemantischen Sphäre der Kommunikation bedeutsam sind, dürfte davon abhängen, wie die zugrundeliegenden Tonhöhenunterschiede bewertet und bei der Interpretation der Aussage vom Hörer herangezogen werden.

Versuche. (1) Ein Teil der ausgemessenen Sätze wurde syntagmenweise mit 5maliger Wiederholung in Hörversuchen angeboten: (1a) einer absolut hörenden Versuchsperson (Angabe von Notenwerten; voll ausgewertete Urteile); (1b) einer Gruppe von 16 Philologiestudenten mit durchschnittlicher Hörerfahrung (Angabe der relativen Tonhöhen durch Striche in einem 3-Liniensystem; 3257 ausgewertete Urteile). Als Übertragungssystem hat dieser Test sendeseitig die Symbole *S* (positive Differenz > 1 Halbton zwischen Iktussilbe und nichtiktischer Silbe), *E* (positive und negative Differenz < 1 Halbton), *F* (negative Differenz > 1 Halbton) und empfangsseitig die Symbole *S'* (positiver Tonhöhenunterschied), *E'* (kein Unterschied), *F'* (negativer Unterschied); Klasseneinteilung bei absoluten Urteilen entsprechend den Festlegungen der Sendeseite (für Gleittöne wurde das Mittel zwischen Anfangs- und Endton genommen), bei komparativen Urteilen nach der Lage der Striche im Liniensystem. Durch Differenzierung nach den verschiedenen Akzenten ergaben sich für (1a) und (1b) jeweils 6 3,3-Matrizen, aus denen folgende Werte (nach Meyer-Eppler) berechnet wurden: zur Einschätzung der Übertragungstreue das Korrespondenzmaß *B* (Spur der Dissipationsmatrix bezogen auf den Symbolvorrat); den informationstheoretischen Wirkungsgrad *WG* als das Verhältnis der



Gi se la May singt em Mon tag in Pa ris

Abb. 1.

richtig übertragenen Information zum Informationsgehalt der gesendeten Symbole; zur Einschätzung der Tendenz der „Falsch“-Bewertungen den Quotienten aus Spaltensumme und Zeilensumme, und zwar $\frac{S'}{S}$ und $\frac{F'}{F}$ (s. Tab. 1). (2) 22 Sätze, die

Tabelle 1.

			Head		Body-Akzente		Nucleus	
B	Einzelhörer	prä	0,75		0,75		0,82	
		post	0,50		0,69		0,81	
	Gruppe	prä	0,50		0,40		0,57	
		post	0,31		0,37		0,41	
WG	Einzelhörer	prä	36 %		40 %		58 %	
		post	7,8 %		27 %		46 %	
	Gruppe	prä	9,2 %		2,2 %		31 %	
		post	1,4 %		1,5 %		5,1 %	
Spaltensumme Zeilensumme	Einzelhörer		S'/S	F'/S	S'/S	F'/F	S'/S	F'/F
		prä	1,12	0,92	0,85	1,63	0,99	1,12
	post	0,97	0,80	0,92	0,87	1,56	0,93	
	Gruppe	prä	1,16	1,12	5,20	0,39	1,67	0,82
		post	0,38	3,80	0,14	6,84	0,77	1,21

in Zuordnungsversuchen 90 —100%ig der Sit. 2 (verbindlich) zugeordnet worden waren, wurden über Tiefpaß ($F_G = 500$ Hz) angeboten, so daß der Stimmklang als Interpretationshilfe ausfiel. Aus den gleichen Sätzen wurden dann alle prä- und postiktischen Silben herausgeschnitten, die kontrastverschärfend wirken konnten. Danach wurden die Sätze wiederum über Tiefpaß vorgeführt (s. Tab. 2).

Tabelle 2.

	Sit. 1	Sit. 2	Sit. 3	Sit. 4	Sit. 5	unklar
ungeschnitten	8,5	42,1	26,1	6,6	10,3	6,4
geschnitten	13,8	38,9	24,0	5,3	11,4	6,6

Ergebnisse. Tab. 1 zeigt: B ist sowohl beim Einzelhörer als auch bei der Gruppe in allen Fällen präiktisch größer als postiktisch. Tonhöhenunterschiede werden also vor Iktussilben eher „richtig“ gehört. Dabei nehmen die „richtigen“ Urteile prä- und postiktisch beim Nucleus den höchsten Wert an. Genauso verhält es sich bei WG. Die Streuinformation bes. bei der Gruppe wird z. T. so stark, daß die Güte der Übertragung auf ein Minimum herabsinkt. Hinsichtlich der Quotienten $\frac{S'}{S}$ und $\frac{F'}{F}$ fällt bei der Gruppe vor allem auf, daß Tonhöhenunterschiede zwischen den Silben, abgesehen von den Stellen, die durch v. Essens bekannte Notierungen festgelegt sind (vor Head, nach Nucleus) in charakteristischer Weise „falsch“ gehört werden, postiktische Differenzen vorwiegend als negativ (bei Body-Akzenten z. B. 7 mal mehr F' -Urteile als angebotene F -Differenzen), präiktische vorwiegend als positiv. Diese Differenzen werden also von musikalisch durchschnittlich veranlagten Menschen selbst bei einiger Übung so unsicher beurteilt, daß kaum eine funktionelle Belastung angenommen werden kann. Diese Annahme wird durch Versuch 2 bestätigt. Tab. 2 zeigt zwar, daß durch die Filterung die Zuordnung zu Sit. 2 unsicherer wird, das Herausschneiden nichtiktischer Silben verändert aber die Streuung nicht wesentlich. Zur Interpretation in der ektosemantischen Sphäre genügt es offensichtlich, daß sich der Hörer bei der Melodiebewertung auf die Iktussilben stützt, auf ihre relative Lage zueinander und auf die Grundtonbewegung in ihnen. Zu untersuchen bliebe, ob solche Differenzen vor allem vor dem Nucleus orthoepisch relevant sind.

LITERATUR

- Elekfi, L.: Untersuchungen zu den Beobachtungsmethoden der ungarischen Sprechmelodie. *ZfPhon.* 18: 9—32 (1965).
- Essen, O. V.: *Grundzüge der hochdeutschen Satzintonation*; 2 Aufl. (Henn, Ratingen/Düsseldorf 1964).
- Maack, A.: Regeln der deutschen Silbenmelodie. *Phonetica* 2: 199—219 (1958).
- Meyer-Eppler, W.: *Grundlagen und Anwendungen der Informationstheorie* (Springer, Berlin-Göttingen - Heidelberg 1959).
- Stock, E.: Untersuchungen zur Intonation bei Stotterern. *Folia phoniat.* 18: 447—461 (1966).

SPEECH INTELLIGIBILITY OF THE DEAF AND HARD OF HEARING PERSONS

I. ŠUPÁČEK

The aim of the reported study was to determine the social effectiveness of the speech in deafness of various degree and duration and to demonstrate the teachers at the Czechoslovak schools for the deaf—where usually much attention is still paid to articulation drill—the extent to which speech intelligibility of the deaf depends upon the accuracy of articulation and upon the musical factors of speech.

Speech intelligibility was investigated in 24 speakers with various hearing losses. Two speakers with normal hearing served as controls. All were 16—19 years old and their I. Q.s revealed normal intelligence. 180 adults, not familiar with the speech of the deaf, served as auditors. In every speaker, recordings of reading 20 sentences, 30 P.B. words and 50 P.B. syllables were made and later scored by 6 male and 6 female auditors.

Fig. 1 shows the results of sentence intelligibility for all groups of speakers both with congenital and acquired hearing losses.

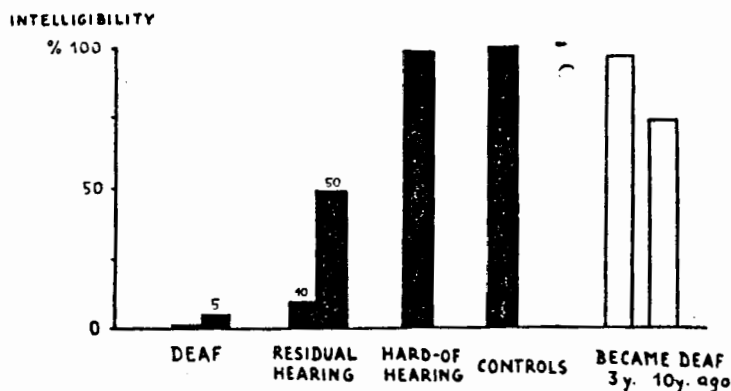


Fig. 1.

Fig. 2 gives the principal result of our investigation. It is the finding that speech intelligibility of the deaf reveals inverse correlations when compared with normal speakers. This can be demonstrated in 3 selected speakers with congenital hearing losses of various degree: The speech intelligibility of a hard of hearing person who could hear loud speech only at a 30 cm distance showed nearly normal results—his

sentence intelligibility being the best = 100%, word intelligibility slightly worse = 90% and syllable intelligibility being the worst = 55%.

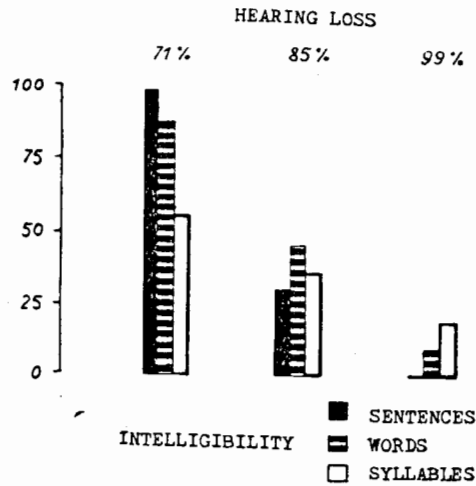


Fig. 2.

Completely reverse results were found with the speech of a congenitally totally deaf—the best intelligibility was found not for sentences but for syllables = 18%, word intel. was worse = 8% and the lowest intel. was for sentences (almost zero).

Between these two cases was the third subject with residual hearing who could only distinguish isolated vowels pronounced closely at the ear; his sentence, word and syllable intelligibility all kept within 30–40%.

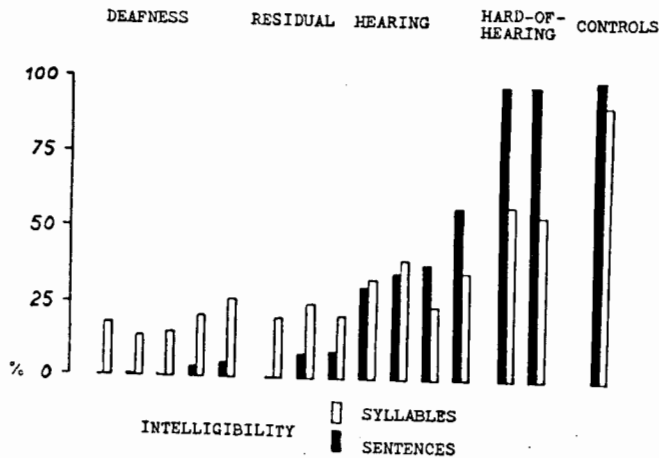


Fig. 3.

Fig. 3 gives the results in all our subjects with congenital hearing losses. Syllable intelligibility equalled 20–50% in all except the controls. Sentence intelligibility,

on the contrary, manifested extreme differences from 0% in the totally deaf to 100% in the hard of hearing persons.

These results show that speech intelligibility of the deaf is relatively best where it mainly depends on articulation, i.e. in isolated syllables. When musical factors play an increasing role in pronouncing word and especially in whole sentences the intelligibility of the deaf speaker rapidly decreases.

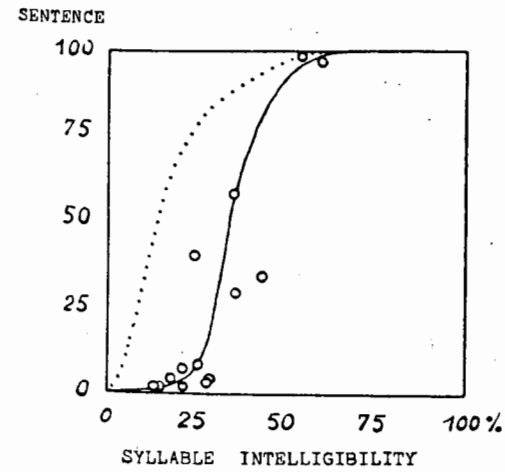


Fig. 4.

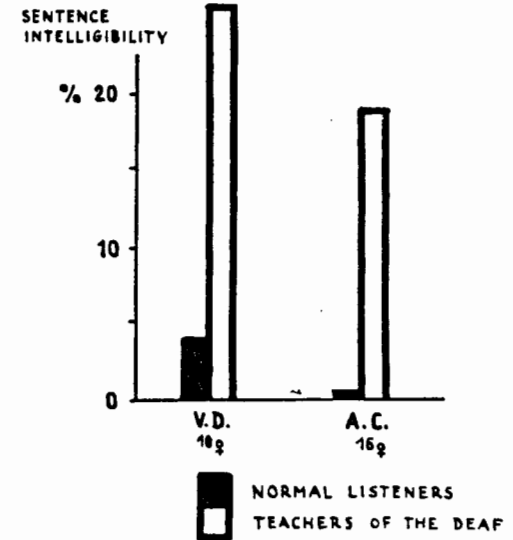


Fig. 5.

In similar foreign investigations as e.g. of Clarke, the listeners were asked to score separately the single musical qualities of the speech. We preferred to use another way of a statistical proof asking the listeners to answer a single and quite a natural question which was very similar to normal situations in communication, that is whether they did or did not understand the deaf speaker. The predominant role of musical factors for the intelligibility of speech was then evident from a simple comparison of the scores for whole sentences, for words and for single syllables.

This relation of syllable and sentence intelligibility is shown in Fig. 4. The full line represents our findings. Compared with the dotted line standing for normal speech it is shifted by approximately 30% to the side of syllable intelligibility. This means that e.g. a deaf speaker with 20% syllable intelligibility has 0 or 10% sentence intelligibility as contrasted with 80% sentence intelligibility of a normal speaker under identical conditions.

Finally we want to mention similar investigations of foreign authors who usually obtained much better results of speech intelligibility. Since these authors used teachers of the deaf as listeners we repeated such an experiment just to demonstrate the difference between scores obtained from untrained and specialist auditors; it

was great, e.g. 1 versus 20 %. It is obvious that specialist auditors cannot give a true picture of the social effectiveness of the speech of the deaf in daily life situations.

Concluding we may say that all the reported results strongly support the necessity of a very early diagnosis and auditory training in all deaf infants since the musical factors are one of the first acquired in every speech development.

SUR LE „LOCUS“ DES CONSONNES ROUMAINES

VALERIU ŞUTEU

En nous proposant d'étudier le locus¹ des consonnes bilabiales, dentales et vélares roumain, nous avons entrepris dans le laboratoire du Centre de recherche phonétique et dialectologique de Bucarest, les recherches nécessaires à l'aide du sonographe.

Les analyses et les observations effectuées sur les données expérimentales, nous permettent de conclure ceci:

Le locus moyen de la consonne bilabiale sourde [p] est de 500 c/s environ, tandis que celui de la sonore correspondante [b] dépasse 700 c/s.

Mais il faut ajouter qu'il y a une grande différence entre la valeur du locus pour [b] initial d'une part, et les valeurs du locus pour [p] initial, [p] final et [b] final, de l'autre.

Si dans le premier cas le locus moyen est de 800 c/s, pour les trois derniers sons mentionnés, le locus moyen est de 400 à 550 c/s.

Il s'ensuit que [b] initial est le son le plus aigu de la série étudiée.

En ce qui concerne les sons [p] initial et [b] final, qui ont une valeur fréquentielle moyenne du locus à 550 c/s, on peut affirmer qu'ils ont un degré moyen d'acuité par rapport à [b] initial et à [p] final.

Le son [p] final est le plus grave (le locus = 400 c/s.)

Les chiffres indiquent donc que [p] est un son plus grave que [b].

Le temps ne nous permet pas d'examiner en détail la situation des consonnes dentales [t] et [d]. Notons seulement le fait que le locus moyen de la sourde [t] se situe à 1550 c/s et que le locus de la sonore [d] est à 1900 c/s environ. Les observations faites pour les bilabiales sont valables pour les dentales.

Il faut insister sur le locus des vélares roumaines, qui ont une situation spéciale.

Les transitions des voyelles [u] et [o] se dirigent vers un point situé à 950 c/s pour [ko], [ku] et à 920 c/s pour [go], [gu]. Quant aux vélares suivies de [a], [ə] et [ɨ], le locus est à 1600 c/s environ pour [k] et à 1750 c/s pour [g]. Enfin, en ce qui concerne le locus des consonnes des groupes [ke] et [ki], il dépasse un peu 2000 c/s.

Pour les vélares finales, également, toutes les transitions ne convergent pas vers le même point.

¹ V. Pierre Delattre, *Les indices acoustiques de la parole*, "Studies in French and Comparative Phonetics", 1966, London, The Hague, Paris, pp. 261—262.

Les transitions de [o], [u] s'entrecoupent dans un point à 800 c/s, les transitions de [a], [ə], [ɨ] s'entrecoupent à 1550 c/s environ et ceux de [e] et de [i] à 2300 c/s environ.

Ces données sont, par conséquent, très proches de celles obtenues pour les vélaires initiales.

Il convient de remarquer que, à la différence des résultats obtenus pour les vélaires du français ou de l'anglais, par exemple, les consonnes vélaires roumaines n'ont pas un ou deux loci, mais trois, du moins pour les vélaires finales; en position initiale les consonnes roumaines des groupes [ke], [ki], [ge], [gi] sont, en fait, palatales.

Cette situation est, à notre avis, un résultat du caractère spécifique des voyelles centrales roumaines [a], [ə] et [ɨ]. Notons qu'en étudiant le locus des vélaires roumaines, nous avons obtenu des données précieuses sur la nature des voyelles mentionnées. Si en français, par exemple, la voyelle [a] fait série commune avec *ɛ*, *e*, *i* (la série palatale) quant au locus des vélaires, le [a] roumain avec [ə] et [ɨ] constituent une série distincte, qui n'est ni palatale, ni vélaire.

Cette constatation confirme l'opinion suivant laquelle le [a] roumain est une voyelle centrale, intermédiaire entre [a] antérieur et [a] postérieur.

Par conséquent, la présence d'un troisième locus pour les vélaires roumaines est parfaitement explicable.

Tandis qu'en d'autres langues nous avons deux loci pour les vélaires, mais une seule série de phonèmes consonantiques postérieurs, en roumain il y a, pour les vélaires, trois loci et deux séries de phonèmes consonantiques (palatales et vélaires).

DISCUSSION

Delattre:

A cette excellente communication je voudrais ajouter que les transitions des spectrogrammes ne sont pas toujours d'accord avec le locus tel qu'il est défini par les expériences de perception. D'après la théorie du locus, une transition ne contribue à la perception d'un certain lieu d'articulation que dans la mesure où cette transition vise (points to) le locus approprié établi par les expériences de perception. Sinon, un autre indice remplace le locus. C'est le cas des vélaires devant voyelles autres que *l*, *e*, *ɛ*, *æ*; l'explosion remplace l'indice du locus.

Fischer-Jørgensen:

Mr. Suteu has found 3 different loci for velars followed by different vowels. I have obtained similar results for Danish velars. Professor Delattre wants to restrict the concept of locus to the meeting point of the transitions found to be relevant for the perception of synthetic speech. I am not quite sure that a locus found by means of stimuli which "the mouth cannot produce" is a useful concept for the description of natural speech; and I do not agree that the transition in *ku* is irrelevant. It indicates that the syllable is not *tu* (but *ku* or *pu*).

Gsell:

Il faut féliciter Monsieur Suteu pour le sérieux et la minutie de son travail. A l'aide d'une méthode classique que nous lui avons montrée à Grenoble il a pu établir que les consonnes vélaires

ont 3 centres de convergence des transitions vocaliques différentes selon qu'il s'agit de voyelles palatales, de voyelles centrales ou de voyelles vélaires. Il y a d'autres méthodes: calcul par ex. des transitions à partir des sections qu'un de nos élèves a développé. Monsieur Delattre a très justement démontré que ces points de convergence des transitions—qui par là définissent la „pente“ des transitions — n'ont rien à faire avec le locus perceptuel, établi par synthèse. On devrait à mon avis dissocier les 2 notions et appeler la première: „locus des transitions“ ou mieux „centre de convergence“. Les 2 notions ne coïncident évidemment pas puisque la parole naturelle (inertie des organes) n'offre pas les conditions idéales de la parole synthétique.

Wittoch:

Au temps de l'évolution historique de la langue roumaine, la consonne „n“ a exercé une influence sur le développement des voyelles précédentes, par ex. sur la voyelle „e“. Est-ce que les recherches de la phonétique expérimentale pourraient contribuer à l'explication de ce phénomène — et en quelle façon?

Suteu:

ad Wittoch: Je ne sais pas. Je n'ai étudié que les consonnes orales. Il faut faire des recherches expérimentales sur les consonnes nasales pour voir si les résultats pourraient donner quelques indications en ce qui concerne la question posée par M. Wittoch.

Je remercie tous de leurs observations utiles et intéressantes.

A STUDY OF RELATIVE EFFICIENCY OF ACOUSTIC PARAMETERS IN THE INTONATIONAL SIGNALS OF AMERICANS ENGLISH

YUKIO TAKEFUTA*

The theory of signal detectability has provided an interesting model for the study of psychophysical phenomena. The difference between the means of the two density functions expressed in terms of their standard deviation is defined as the index of signal detectability in this theory. Utilizing this index, one can quantify relative effectiveness of any measurable acoustic parameters as elements of certain phonetic or linguistic signals by plotting a frequency distribution of "the stimuli with a phonetic signal to be studied" and another distribution of "the stimuli without such a signal" on a scale of the measured acoustic parameter. In this study, the indices have been calculated to assess the relative effectiveness of 27 arbitrarily chosen measures of five acoustic parameters as possible elements of the phonetic signal of American intonation.

Sixteen native speakers of American English listened to 400 pairs of English sentences recorded by 20 native American speakers, and reported whether the paired sentences were read with the same or different intonations. The difference in the intonations perceived by native listeners between the paired sentences was considered the phonetic signal of the intonation in this study. Specifically, a pair of sentences which was heard by 12 or more of the 16 listeners to be different in intonations was treated as "a stimulus with an intonational signal." Conversely, a pair of sentences which was heard by 12 or more of the listeners to be the same was treated as "a stimulus without an intonational signal." From the initial 400 stimuli, two sets of 94 stimuli, the "SET SN" having an intonational signal and the "SET N" not having such a signal, were randomly selected and subjected to acoustic analyses of (1) pitch pattern, (2) pitch level, (3) intensity, (4) rate, and (5) the extent of pitch change. A measure of the signal strength of each stimulus in an acoustic parameter was made by calculating the average of the differences of the measure obtained by comparing the paired sentences syllable by syllable.¹

Several problems had to be resolved in the process of data reduction. The two

*The Ohio State University, Columbus, Ohio, U.S.A.

¹ Syllables were defined in this study as the equal time segments obtained by dividing the time spent in reading the total sentence by the number of the actual syllables defined phonetically.

rules proposed by Witting² (two "dimensional rules" concerning duration and frequency) and the data on the threshold of pitch modulation³ expressed in terms of the direction, duration, and the extent of frequency modulation were used to simplify and schematize the pitch pattern. Quantification of the schematized pitch pattern was made by assigning a number 3, 2, and 1 to a rising, level, and falling pattern respectively.

Data reduction with regard to the measurement of the pitch level and the intensity was attempted from two directions: first, by sampling only one representative value

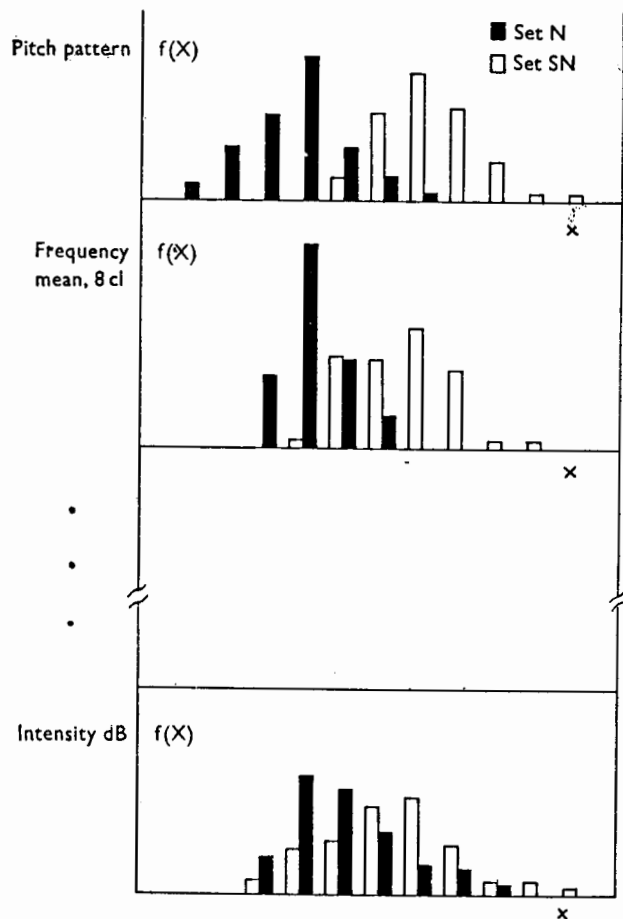


Fig. 1. Frequency distributions of the two sets of stimuli: SET SN and SET N, plotted on the scale of each physical measurement.

either by measuring the center of each syllable or by taking the average of all the raw data of the syllable; second, by making the units of contrastive scale more gross by dividing the total range of pitch or intensity of one speaker into equal contrastive scale units of 8, 4, or 2. The second method of data reduction was used also for the reduction of the data of duration and the extent of frequency change. By testing several different methods of data reduction in measuring each acoustic parameter 27 different physical measures of the signal strengths were obtained of the two sets of stimuli: SET SN and SET N. The frequency distributions (histograms) of the signal strengths for the SET SN and for the SET N were plotted on the scale of each of these 27 physical measurements (Figure 1). Assuming the normality in these distributions, 27 indices of signal detectability, in turn, were obtained from the information of the difference in the means and the standard deviation of the paired distributions. The indices varied from 0.1 to 2.1 (Table 1). No individual measures were as good as the average signal detectability calculated from the signal detection of an individual native American listener.⁴ However, the combined infor-

Parameter	Fundamental Frequency (mean)				Fundamental Frequency (center)					
Scale	Hz	8 cu	4 cu	2 cu	Hz	8 cu	4 cu	2 cu		
Index of Detectability	1.1	1.4	1.2	0.7	0.8	1.1	1.1	0.9		
Parameter	Extent of Frequency Change				Rate of Speaking					
Scale	Hz	8 cu	4 cu	2 cu	msec.	8 cu	4 cu	2 cu		
Index of Detectability	0.5	0.7	0.6	0.6	0.8	0.4	0.1	0.2		
Parameter	Intensity (mean)				Intensity (center)					
Scale	Ratio	8 cu	4 cu	2 cu	dB	Ratio	8 cu	4 cu	2 cu	dB
Index of Detectability	0.3	0.2	0.1	0.1	0.5	0.7	0.2	0.1	0.1	0.2
Parameter	Pitch Pattern				Pitch Pattern and Fundamental Frequency					
Scale	3 cu				3 cu and 8 cu					
Index of Detectability	1.2				2.5					

Table 1. Summary of 28 indices of signal detectability obtained from 27 measures of five physical parameters and from a combined data of pitch pattern and pitch level (CU stands for the contrastive units).

² C. Witting, "A Method of Evaluating Listeners Transcriptions of Intonation on the Basis of Instrumental Data," *Language and Speech*, 5, 1962, 138-150.

³ Y. Takefuta, "Perception of Frequency Modulation," Paper prepared for the 74th Meeting of the Acoustical Society of America, Miami Beach, Florida, U.S.A., 1967.

⁴ Y. Takefuta, "A Study of Relative Efficiency of Acoustic Parameters in the Intonational Signal of American English," unpublished Ph. D. dissertation, The Ohio State University, Columbus, Ohio, U.S.A., 1966. The average signal detectability of a native American listener in the detection of an intonational signal in American English was 2.4.

mation of the most efficient physical measure (pitch pattern) and the second most efficient measure (pitch level) elicited a detectability index of 2.5 which was about the same as the signal detectability of an average native listener.

From the results of the present study, it was found that the order of relative efficiency of five acoustic parameters as possible intonational signals were (1) pitch pattern, (2) pitch level, (3) extent of pitch change, (4) rate of speaking, and (5) intensity. Efficiency of any acoustic parameter changed considerably by using different methods or units of measurement. The combined information of pitch pattern schematized according to the rules based on the findings of psychophysical experiments and the pitch level measured in eight contrastive units seems to comprise the primary phonetic signal of intonation, perceived by average American listeners.

REFERENCES

- Cowan, J. M. "Graphical Representation of Perceived Pitch in Speech," *Proc. 4th Int. Cong. Phon. Sci.*, Helsinki, 1961, 567—570.
- Sidowski, J. B. (ed.) *Experimental Methods and Instrumentation in Psychology*. New York: McGraw-Hill, 1966.
- Swets, J. A. (ed.) *Signal Detection and Recognition by Human Observers*. New York: John Wiley & Sons, 1964.
- Takefuta, Y. "A Study of Relative Efficiency of Acoustic Parameters in the Intonational Signal of American English." Unpublished Ph.D. dissertation, The Ohio State University, Columbus, Ohio, U.S.A., 1966.
- Takefuta, Y. "Perception of Frequency Modulation" (Paper prepared for the 74th Meeting of the Acoustical Society of America, Miami Beach, Florida, U.S.A., 1967).
- Witting, C. "A Method of Evaluating Listeners Transcriptions of Intonation on the Basis of Instrumental Data," *Language and Speech*, 5, 1962, 138—150.

ACKNOWLEDGMENT

This research was supported by a contract between the Office of Naval Research and The Ohio State University Research Foundation, Nonr 495(18).

DISCUSSION

Black:

These are very interesting results. I would appreciate your telling us how the interspeaker differences in pitch were controlled.

Takefuta:

ad Black: The interspeaker differences in pitch were controlled by subtracting the median pitch of each speaker from all the measurements before comparing them.

ЗВУКОВАЯ ФАКТУРА СТИХА И ЕЕ ВОСПРИЯТИЕ

К. Ф. ТАРАНОВСКИЙ*

В своем докладе я ограничусь только одним вопросом: о роли ударных гласных в русском стихе и о синэстетическом их восприятии.носителем информации, воспринимаемой вне когнитивного плана, в русских ударных гласных являются два различительных признака: низкая тональность / высокая тональность и диффузность / компактность. Низкая *первая форманта* характеризует диффузные, а высокая — компактные гласные. Низкая *вторая форманта* определяет низкотональные, а высокая — высокотональные. Низкотональные *у* и *о* противопоставляются высокотональным *и* и *э* как *темные* — *светлым*, и в прямом значении (мрак — свет) и в переносных значениях (мрачное настроение — радостное или умиротворенное настроение, отрицательное явление — положительное явление и т. п.). Гласные *ы* и *а* занимают среднее положение между *у* — *о* и *э* — *и* и в „световом“ противопоставлении, в общем, не участвуют. Правда, в окружении темных *у* и *о* гласный *а* легко примыкает к ним, как напр. в стихотворении „Свет“ З. Гиппиус: „(1) Стоны, / Стоны, / Истомные, бездонные, / Долгие, долгие звоны / Похоронные, / Стоны, / Стоны . . . // (2) Жалобы, / Жалобы на отца . . . / Жалость язвущая, жаркая, / Жажда конца, / Жалобы, / Жалобы . . . // (3) Узел туже, туже, / Путь все круче, круче, / Все уже, уже, уже, / Угрюмей тучи, / Ужас душу рушит, / Узел души, / Узел туже, туже, туже . . .“ — В этих трех строфах ударные гласные распределяются следующим образом: в первой 10 *о*, во второй 10 *а*, а в третьей целых 20 *у*. Темная звукопись явно вторит содержанию текста. В четвертой строфе высокотональные гласные резко контрастируют с низкотональными: „Господи, Господи — нет! / *Веще сердце верит!* / Боже мой, нет! / Мы под крылами твоими. / Ужас. И стоны. И тьма . . . а над ними / Твой *немеркнувший свет*.“ — Ключевыми словами в этой строфе являются *верить* и *немеркнувший свет*. — Из десяти ударных *а* второй строфы девять являются задним, более темным вариантом фонемы /*a*/; этот вариант легко включается в низкую тональность темных *о* и *у*. Более светлый передний вариант фонемы /*a*/, в свою

* Гарвардский университет, США.

очередь, не контрастирует с высокотональными гласными, как, напр., в „Последней любви“ Тютчева: „Сияй, сияй, прощальный свет, / Любви последней, зари вечерней...// Помедли, помедли, вечерний день, / Продлись, продлись, очарованье.“ — В статистических анализах темных и светлых гласных русский гласный *a* следует учитывать как нейтральный.

Полюсом компактности является гласный *a*, наделенный самой высокой первой формантой. *A* противопоставляется диффузным *y*, *ы* и *и*, т. е. гласным с самими низкими первыми формантами. Фонемы /э/ и /о/ в этом отношении следует признать нейтральными, хотя их „открытые“ варианты примыкают к компактному *a*. Синэстетически компактность ассоциируется с чувством объемности, полноты, величия, равновесия, силы, т. е. с переживаниями, которые условно можно объединить в понятие *устойчивости*. Диффузность, наоборот, ассоциируется с *неустойчивостью*, с отсутствием равновесия, с душевным разладом. На диффузных *y*, *ы*, *и* построена первая строфа „Молитвы“ Лермонтова: „В минуту жизни трудную / Теснится ль в сердце грусть: / Одну молитву чудную / Твержу я наизусть.“ Ключем к этой звукописи является *трудная минута жизни*. Еще Ломоносов писал в „Кратком руководстве к красноречию“ (§ 172), что „частое повторение писмени *a* способствовать может к изображению великоления, великого пространства...“ — Чувством объемности несомненно проникся и Блок, когда написал строки: „О весна, без конца и без краю — / Без конца и без краю мечта.“ Блок, пожалуй, чаще других поэтов насыщал компактным *a* свои строки. Два ударных *a* содержатся и в самом имени „Прекрасная Дама“, и ее образ часто сопровождается ассонансами на *a*: „Она молода и прекрасна была...// Она росла за дальними горами...// Она стройна и высока...// Ты в поля отошла без возврата...“

Если компактность связана с нейтральным гласным *a*, в световом противопоставлении не участвующим, то диффузность сочетается и с темным *y* и с светлым *и*. В гласном *y* диффузность не противоречит низкой тональности, ибо и та и другая легко включаются в отрицательное семантическое поле, как, напр., в выше приведенных строках З. Гиппиус: „Ужас душу рушит, / Узел души, / Узел туже, туже, туже...“ Неясно, могут ли в поэтических текстах ассонансы на *и* одновременно функционировать и как светлые, и как диффузные, напр., одновременно отражать и восторг и душевный разлад. Характерных примеров такого сочетания у меня нет. При некотором воображении можно почувствовать и свет и душевную боль в тютчевской строке: „Продлись, продлись, очарованье“. В современной жизни такое сочетание наблюдается: американские девушки приветствуют своих любимцев-певцов визгом „И-и-и“, несомненно выражающим истерический восторг.

Изучение поэтических текстов и теоретических высказываний поэтов дает нам надежный материал для суждений о синэстетическом восприятии

звучков. Результаты этого изучения можно проверять и при помощи „метода придуманных слов“. В умной и интересной статье Е. В. Орловой „О восприятии звуков“ (*Развитие фонетики современного русского языка*, Москва, 1966) описан эксперимент, построенный на световом противопоставлении *темный — светлый*, отождествленном с эмоциональным противопоставлением *плохой — хороший*. Фонема /*a*/ расценивалась автором как темная, а компактность и диффузность вовсе не были учтены. Некоторые результаты эксперимента оказались для экспериментатора неожиданными. Предполагалось, что слово *Агабák* входит в отрицательное семантическое поле. Между тем, большинство информантов дало ему положительную оценку. Напр.: „Юноша 20 лет, высокий, красивый, хорошо сложен. Очень волевой, сильный“. Очевидно информант почувствовал в слове *Агабák* компактность пяти фонем (3 /*a*/, /*г*/ и /*к*/). Слово *Терíлья* получило противоречивые оценки. Положительная оценка основывалась на высокой тональности первых шести звуков: „Красивая, умная, нежная, благородная, впечатлительная“. Отрицательные оценки основывались на диффузности фонемы /*и*/: жадный, скупой; хитрая, злая, плохая; фашист; изменница; предатель. Метод придуманных слов имеет еще один недостаток. Экспериментаторам не удается полностью избежать непредвиденных семантических ассоциаций. Напр., один информант воспринял слово *Терилья* как испанское имя: „Испанка лет 30.“ Все же и он почувствовал в этом слове диффузность, прибавив: „Страстный, необузданный нрав“. Нам кажется, что изучение поэзии дает более надежный материал для суждения о синэстетическом восприятии звуков, чем метод придуманных слов.

(См. также мою статью „The Sound Texture of Russian Verse in the Light of Phonemic Distinctive Features“, *IJSLP*, IX (1965).

DISCUSSION

Напп:

Из гласных какие, — в рамках сопоставлений бинаристических — Вы считаете признаковыми (темные или светлые), а какие беспризнаковыми?

Where poets and poetry are concerned I would like to call the attention to the fact, that acoustic structures often do induce them to invent words—or to take them from distant (own or foreign) dialects.

Taranovsky:

Ad *Напп:*

Это, пожалуй, самый сложный вопрос фонологии. Я склоняюсь считать низкотональные (темные) гласные маркированными.

Yes, you are right. For example, Russian trans-sense poetry should definitely be analyzed in terms of phonemic distinctive features. But the morphophonemic level must also be taken into account: affixes assign invented words to certain semantic fields, possibly, implying +/— quality. Studying poets' manuscripts is very helpful as well. Changes of the text are very often motivated by the sound texture of the words involved.

UNUSUAL MISTAKES IN HEARING AND PRONOUNCING FOREIGN SOUND-SEQUENCES

ANA TĂTARU

When arnileng the pronunciation of a foreign language adults often fail to utter the right sounds and sound-sequences because of their own speech habits which are deeply fixed.

We have considered that, apart from such usual mistakes due either to native or to some assimilated foreign articulatory habits, there may also occur some unusual hearing and pronunciation mistakes due to the influence of the memory of certain sound-sequences representing words acoustically similar to the foreign words. The present paper attempts to prove that.

In teaching English to adults during the last few years, we have met among our Romanian students some of whom, along with their mistakes typical of Romanian speakers, also make mistakes of a French type. As they learned French at school for a few years only, it could not be a matter of French sounds having penetrated deeply into their native articulatory habits. These students have some important characteristic features in common:

1. as to their backgrounds, all of them are peasants' children, born and brought up in purely Romanian villages of Transylvania, and they had not even heard any other language spoken there. Their acquaintance with foreign languages was acquired at school, where they learned some French and very little Russian. From the village public school they came to the university and, apart from their main subject, they took English for their three-year foreign language courses.

2. During this period of time, they have learned to master English pronunciation with greater difficulty than other Romanian students.

3. Their French-type mistakes never appear when they practice the pronunciation of isolated English sounds; they only appear in sound-sequences.

4. Their substitution of French sounds for English sounds does not occur in all sequences of the same kind, but only in those with such acoustic forms as can remind them of previously learned French words. For instance, they utter correctly the personal pronoun *you*, but, incorrectly the word *view*, pronounced either [vjy, vy] or [vjø] depending on the French word [vy] (vue) or [vjø] (vieux) they have in mind.

5. Their mistakes have nothing to do with the French reading rules.

6. Their French-type mistakes occur in English new words they are learning, as

well as in words already learned if these disappear for some time from the given exercises and then suddenly reappear.

The exercises we have used to convince ourselves of the real cause of this kind of mistake have been speech exercises with words in isolation or in sentences, the students having either to imitate the teacher's pronunciation—the written form being present or absent—or to try first to utter those words with the written form before them and afterwards to listen to the teacher's pronunciation. (On the accompanying tape we recorded the pronunciation of two beginning students, a girl and a boy, the former during the first term of the first year, the latter during the first term of the third year. The text on the tape includes their French-type mistakes in pronouncing the English [ju:] sequence of certain words, as well as in pronouncing some other words to prove that their mistakes are not influenced by French reading rules.)

In conclusion, we think that the variety of mistakes made by adults learning the pronunciation of a foreign language may also include the mistakes here mentioned in which the memory of the acoustic forms of certain words (those belonging to a previously imperfectly learned foreign language) plays the main part. At the same time we believe that this kind of mistake is more possible and is fully verifiable with speakers who have not in mind the acoustic units of several foreign languages. In this way their mind is still very receptive and the acoustic images not having overlapped, are clear and strong enough to exercise their influence on occasion.

DISCUSSION

Martens:

When students learn a second foreign language one must always bear in mind that the student will subconsciously be somehow guided by his knowledge of structure of the first foreign language he has learnt. A student ought to be without reading knowledge, if we want to know whether a pronunciation of Engl. "view" as French "vue" is not subconsciously guided by a faint resemblance of the spelling.

SPECIFYING SPEECH

M. A. A. TATHAM*

The present paper is concerned with problems arising from the specification of the speech act. It is intended that the production model of which this is a preliminary consideration should reflect the human process of generating meaningful speech from processed linguistic units. The concern therefore is not with suggesting any static correlation between *a sound* and linguistic units but with a dynamic specification for generating the one from the other.

It has been pointed out (Ladefoged 1967) that Jakobson, Fant and Halle in their Distinctive Feature Theory (JHF 1951) did not introduce any combinatory restrictions on their features. This enables both impossible and possible combinations to be generated, including among those possible many that have not been found among the world's languages. No doubt this state of affairs arises from the original descriptive purpose of the theory where the combinatory restrictions are implicit in the sounds being described.

The present production model proceeds in terms of a featural matrix based on those articulatory or neuro-physiological parameters found not to be "significant" or "distinctive" in the languages of the world, but *necessary for modelling the available data about production*. Thus in this case interacting rules for featural combination must be made available on two axes: vertical and horizontal.

The rules for generating vertical combinations of the features may be considered on three major levels: (a) those rules excluding physiologically impossible combinations, (b) rules restricting the possible combinations to those actually found in the world's languages and (c) language specific rules which determine which of the combinations arrived at under (b) can be used in any one language or dialect.

As Ladefoged has underlined (1967) it is necessary at the level of physical phonetics to be able to compare not only sounds as they occur and contrast within one language but also across languages. The best way I can see of doing this is to provide a means of generating an inventory of sounds from a restricted set of features and then limit this by providing a set of language specific rules drawing upon this inventory.

* University of Essex.

It has been proposed by several researchers (Holmes et al. 1964; Öhman, 1966, 1967) that one way of approaching the mapping of linguistic units into the continuous event of running speech is to regard the speech act as a series of linked, missed targets. This usually assumes some phonemic programming of the articulators, modified by rules reflecting the physical restrictions imposed upon their linear combination: the so-called coarticulation rules. This approach has the advantage of overcoming the previously outstanding problem of segmentation.

Dr Fromkin's work (1966) in electromyography of the lip muscles tends to indicate that certainly for this parameter it is not simply a question of missed targets and coarticulation phenomena. The action potential resulting from *same* linguistic phonemes is not identical in different contexts (particularly where the same phoneme occurs in initial and final position). Assuming that the EMG signal is directly related to the neural impulses controlling the muscle this would imply that the differences in this case may result from certain positional restrictions imposed at a higher level; this ties in very conveniently with the present proposal. Unfortunately, extensive EMG work at Leeds University Phonetics Laboratory (Tatham and Morton forthcoming), repeating Dr Fromkin's experiments and extending her field, has not produced similar results, at least not for the duration and amplitude of the EMG signal from the lip muscle. Those differences which did show were not decisive enough to support Dr Fromkin's hypothesis completely.

However, it is obviously true that there *are* positional variants of phonemes, although some of these (eg. initial and final (t) in English) are not variants of necessity. Whether these are a direct result of the kind of programming of the articulators that Dr Fromkin suggests or whether they are the result of a cadence-bound rearrangement of priorities among articulatory parameters remains to be seen. It might well be that some simple phonemic programming is overridden at a high level under certain conditions by prosodic features and that syllable or larger units must be accounted for.

The division of the model along two axes: (a) mapping of linguistic phonemic units into extrinsic allophones and then into intrinsic allophones (Ladefoged 1967) and (b) permitting this axis to draw at the central point upon a range of extrinsic allophones derived from restriction on featural combinations produces a more meaningful and powerful model when higher level combinatory restrictions are allowed for on the featural level. Neuro-physiological features should be mapped individually, but in an explicitly related fashion, thus permitting a less clumsy representation of the available data.

REFERENCES

- Fromkin, Victoria A., Neuro-Muscular Specification of Linguistic Units, *Language and Speech*, 9, 170, 1966.
- Holmes, J. N., I. G. Mattingly and J. N. Shearme. *Language and Speech*, 7, 127, 1964.
- Jakobson, R., G. Fant and M. Halle, *Preliminaries to Speech Analysis*, MIT Press, Cambridge, Mass., 1951
- Ladefoged, P., *Linguistic Phonetics*, Working Papers in Phonetics 6, UCLA, 1967.
- Öhman, S.E.G., Coarticulation in VCV Utterances: Spectrographic Measurements, *JASA*, 39, 151, 1966;
- Numerical Model of Coarticulation, *JASA*, 41, 310, 1967.
- Tatham, M. A. A. and Katherine Morton, forthcoming paper documenting recent EMG experiments.

PERCEPTION OF LOW-REDUNDANCY SPEECH IN CHILDREN

G. P. TEATINI*

Our experiment was performed upon 80 children attending the first four classes of a primary school.

20 subjects were tested in every class. Their age ranged from 7 (1st class) to ten years (4th class).

The children were separately tested with speech messages, which had been recorded on a tape and delivered through ear-phones at a comfortable intensity level (about 50 dB over threshold). Speech material consisted of 9 lists of 10 sentences each: in the first list there was "normal" redundancy speech, whilst in the remaining 8 lists there was low-redundancy speech: 4 lists were composed of distorted sentences and 4 of time-compressed sentences.

Sentences were plain declarative utterances: their length ranged from a minimum of 8 to a maximum of 16 syllables, with a modal length of ten syllables.

The first list of "normal" redundancy sentences was delivered with the purpose of making the children well acquainted with their task. Each subject repeated the sentences of this list without errors.

Then followed the distorted and time-compressed sentences.

The intelligibility of the speech was measured, after the following criteria: the repetitions without errors were judged correct as well as the responses with some minor mistakes (usually represented by the exchange of singular for plural).

Every other kind of response was judged incorrect.

Results are reported in figures. The percentage of sentences correctly repeated is shown on the ordinate; the school grade is pointed out on the abscissa. The dotted area around the curves is the standard deviation.

This first figure refers to distorted sentences. In the higher part of the graph is shown the significance of the comparisons after Student's test. We see that intelligibility increases from the first to the 3rd class with statistically significant differences; on the contrary the scores in the 4th class are quite close to the 3rd one.

Time-compressed speech shows a remarkable increase of intelligibility when passing from the 1st to the 2nd class; then the growths become small and non significant.

* From the Otolaryngological Clinic of the University of Ferrara, Italy. Head: Prof. E. Bocca.

On the whole, it seems that the functions have an asymptotic course: as a matter of fact, we may extrapolate the curves by making use of the scores which have been determined in adults with similar speech material. The percentage of sentences correctly repeated by a group of subjects 20—30 years old with a primary school degree is practically the same as the one of our 4th class children.

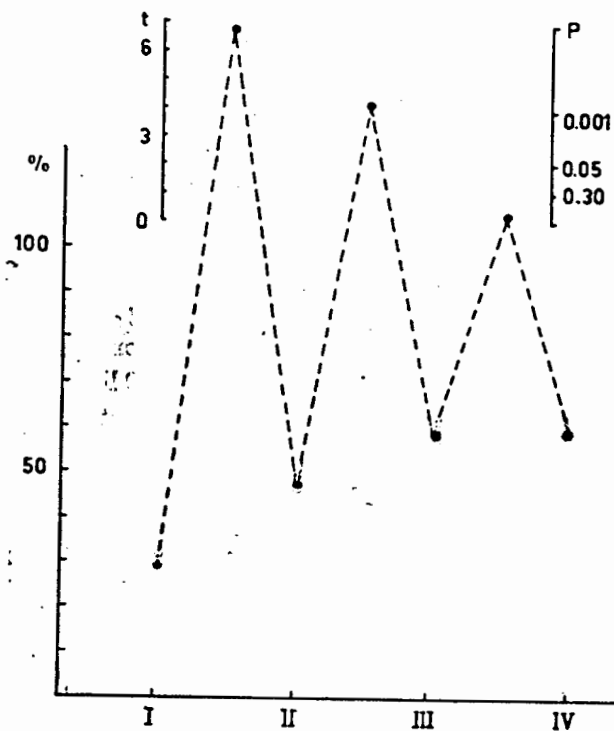


Fig. 1.

The incorrect responses may be divided into three categories:

a) the first group includes repetitions with errors of grammar. They were rather uncommon (only the 8% of all incorrect responses) and more than $\frac{3}{4}$ of them were obtained from children of the first two classes.

b) the second group includes repetitions which were grammatically correct but contained mistakes of semantics. Such a kind of errors occurred in the 36%; and again more than the half were made by 1° and 2° grade children.

c) in the third group, which numbers 56% of all incorrect responses, the sentences were not repeated at all. This happened more frequently in the older subjects.

Let us now draw some conclusions from these results.

The operational model of our experience may be concerned as triad of functions: input-transmission, central processes, output-transmissions.

At their turn, the central processes may be shared in three stages: a) discrimination b) integration c) symbolization-comprehension.

When we deal with "normal" redundancy speech, we may get a correct repetition, i.e. a good intelligibility provided that the single phonemes are integrated. An instance of repetition without comprehension is that stage of language development which was defined "echoic", at the beginning of the second year of life.

On the contrary, when we are concerned with low-redundancy speech (as it is the case of our distorted or time-compressed sentences) we may obtain a correct repetition only if the single forms of the utterance are identified and understood: the identification may require the interpolation of those phonemes which have been missed owing to the distortion.

In order to get a correct repetition it is not indispensable that the contextual meaning of the sentence is understood; but if the meaning too is understood, the correction of the mistakes which are due to the distortion is no doubt made easier.

Since the faculties of comprehension and symbolization are functions of language learning, it follows that the intelligibility of low-redundancy speech increases with increasing age; the greatest increments are seen in the first school years and became gradually smaller as the age grows, until they reach an asymptote.

DISCUSSION

Bystrzanowska:

We have made similar observations among older people listening to the speech in noise. In normal conditions their hearing in noise is the same as among young people. When they are tired their discrimination ability is lowered by the noise more than in young people.

MOTIVATION UND INFORMATION BEI DER LAUTGEBUNG VON SÄUGETIEREN

G. TEMBROCK*

In ethologischer Sicht beruht die Transformation bestimmter Informationen zu einer Nachricht beim Expedienten auf einem Ritualisationsvorgang. Dadurch werden Gebrauchssysteme in Signalsysteme umgewandelt. Voraussetzung ist dabei die Mitwirkung eines potentiellen Rezipienten. Die Eignung zur Ritualisation wird dabei bestimmt durch den Grad der Übereinstimmung von Epiphänomenen, mit denen sie zugrundeliegende neurale Vorgänge abbilden. Erregungssysteme arbeiten entweder explosiv oder graduell; bei entsprechender Koinzidenz können die Begleitprozesse entsprechend diskret oder analog codieren. Je präziser in solchen Epiphänomenen (z. B. hörbares Atmen) bereits Prinzipien, nach denen Informationen im Nervensystem codiert sind, wirksam werden (potentielle Information), desto leichter können sie von einem möglichen Empfänger, der vielleicht die spezielle Systemlage des Expedienten auslöste, auch entschlüsselt werden (aktuelle Information).

Für die Evolution der Signalsysteme ist erforderlich, daß durch die Signalübertragung vom Expedienten auf den Rezipienten, bestimmte Steuer- und Regelungsvorgänge verbessert und die optimalen Fließgleichgewichte eingestellt werden.

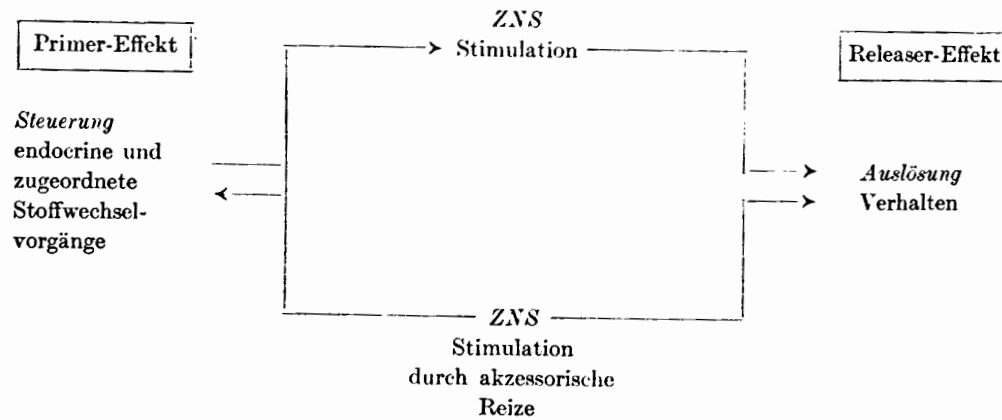
Da die Evolution diesen Ritualisationsvorgang nur schrittweise vollziehen kann, ist die Frage zu stellen, über welche primären Wechselbeziehungen zwischen Sender und Empfänger sich der akustische Kanal der Informationsübertragung überhaupt aufbauen konnte.

Dafür vermuten wir als Ausgangssituation jene Funktionsbereiche, in denen innerhalb einer Population als dem eigentlichen Evolutionssystem die Genverteilung unmittelbar bestimmt wird: die Fortpflanzung. Diese erfordert (unter Einschluß der Jungenaufzucht) bei den Säugetieren mindestens eine vorübergehende Kommunikation zwischen wenigstens zwei Individuen (bei Jungenaufzucht meist mehr). Die entscheidende Primärleistung ist hier die Distanzregulation. Die zwischentierliche Signalübertragung stellt Sollwerte des interindividuellen Abstandes ein. Diese können, je nach Systemlage, auch zwischen zwei bestimmten Individuen einer Population sehr wechselnd sein. Als prinzipielle Möglichkeiten ergeben sich damit:

Distanzerhaltung → Distanzverminderung (affin)
 → Distanzvergrößerung (diffug)

* Zoologisches Institut der Humboldt—Universität, Berlin.

Da bei Säugetieren in der Stammesgeschichte anfangs der chemische Kanal dominierend war (die Reptilien hielten den optischen besetzt), könnte sich der akustische aus Redundanzen entwickelt haben, also auf bereits bestehenden distanzregulierenden Systemen. Dabei lassen sich zwei Wirkungsprinzipien am Rezipienten unterscheiden: der Primer- und der Releaser-Effekt (vergl. Wilson u. Bossert):



Das affine System dient einmal der Regulierung einer Nahdistanz, zum anderen der Verminderung einer Ferndistanz. Die zugeordneten Laute können dementsprechend als Kontaklaute und Distanzlaute (Ruflaute) bezeichnet werden. Die erste Gruppe wird vielfach als „Stimmföhlung“ bezeichnet. Beide Lautgruppen zeigen eine Tendenz zur Stimmhaftigkeit, die bei den Säugetieren sicher damit zusammenhängt, daß durch das Frequenzspektrum spezielle Informationen codiert werden können. Dieses aber wird gerade bei den Distanzlauten von besonderer Bedeutung und muß daher bereits bei den Kontaktlauten eingeleitet werden. Für die Kontaklaute ist in erster Linie die Information über Artzugehörigkeit von Bedeutung, falls sie nicht nur redundant sind, Kontaklaute sind „ungerichtet“ (also nicht „adressiert“). Es sind kurze, ein- bis zweisilbige Laute, oft wiederholt und relativ leise, da sie geringere Distanzen einregeln. Die Wiederholung könnte eine zusätzliche Information über die Raumlage (Ortung) enthalten. Außerdem „neutralisieren“ derartige Laute potentielle diffuse Wirkungen, die durch die Nähe eines Artgenossen ausgelöst werden können. Das ist besonders wichtig, wenn der Sichtkontakt häufig unterbrochen ist (dichte Vegetation). — Wiederholung schließt die Möglichkeit leichter Variabilität des Grundlautes ein, weil dies Adaptationen verhindert und die Ortung erleichtert. Diese wird auch durch ein breiteres Frequenzspektrum gefördert.

Aus diesen Lauten haben sich dann die Distanzlaute entwickelt, und zwar die affinen distanzvermindernden Laute, die vor allem dann funktionell werden, wenn andere Kommunikationsmittel versagen. Hier geht der Faktor Zeit in das System mit ein als Spanne zwischen der Schallabstrahlung und dem ersten feedback. Distanzverminderung erfordert außerdem Überwindung ethologischer Schranken und setzt

oft eine „Umstimmung“ im Empfänger voraus, so daß diese Laute eine Primer-Wirkung haben können. Die Effektivität dieser Lautäußerungen wird umso größer, je mehr Informationen sie verschlüsseln, Artzugehörigkeit, Alter, Geschlecht, Individualstatus, physiologischen Status und vielleicht auch den Umgebungsstatus betreffend. Die Wahrscheinlichkeit der Wirksamkeit der Distanzlaute ist mit dem mittleren Informationsgehalt korreliert. Der Faktor Zeit gibt die Möglichkeit, anstelle eines Parallelsystems (auf das durch das Frequenzspektrum nicht ganz verzichtet wird) ein Seriensystem einzusetzen, bei dem aufeinander folgende Signale die Nachricht codieren. Dies kann geschehen durch: Lautdehnung, Nuancierung, Strophenbildung, Frequenzmodulation, Einbau einzelner Laute (damit Bildung heterotyper Rufreihen bis zu „Gesängen“).

Das distanzvergrößernde System nennen wir diffug. Hier stellt der Rezipient eine Störgröße dar, hier steht Genauigkeit gegen Kürze; die Lautäußerung ist dann optimal, wenn sie Releaser-Wirkung hat und sofort den Rückzug des Rezipienten hervorruft. Dazu gehören ein möglichst starker Reizkontrast zu den affinen Lauten sowie Eigenschaften, die Adaptationen verhindern. Solche Defensivlaute haben Geräuschcharakter, sind kurz und werden explosiv geäußert, wenn sie voll ausgebildet sind. Auch hier können zwei Gruppen unterschieden werden:

- a) Diffuge Kontaklaute, im agonistischen Verhalten vom unterlegenen Tier geäußert.
- b) Diffuge Distanzlaute, bei Unterschreitung einer kritischen Distanz durch den Rezipienten.

Die Systemlage erfordert bei den diffugen Kontaktlauten häufige Wiederholung, aber unregelmäßig, um Adaptationen zu vermeiden, während die diffugen Distanzlaute in größeren Abständen oder nur einmal geäußert werden.

Eine Sonderstellung nehmen die „Warngeräusche“ ein, die wahrscheinlich primär dem Ruheverhalten zugeordnet sind und sich durch Ritualisation emanzipiert haben. Sie sind wieder arbiträr und haben die Funktion, die Unterschreitung einer bestimmten Distanz zu verhindern. Dementsprechend sind sie langgezogen und werden solange geäußert, bis der Effekt erreicht ist, anderenfalls folgen diffuse Distanzlaute. Hierher gehören Geräusche wie das Zischen, Fauchen und Knurren.

Bestimmten Motivationslagen sind schließlich ambivalente Lautformen zugeordnet, die sowohl affine wie auch diffuse (oder warnende) Elemente enthalten können, wie beispielsweise Laute der Territorialanzeige, die Geschlechtspartner anlocken, Rivalen aber abschrecken. Hier sind zwei Prinzipien verwirklicht:

- a) Parallelsysteme (Mischlaute): Überlagerung zweier Lautformen,
- b) Seriensysteme: Zeitmuster. Strophen, in denen verschiedene Lautformen aufeinanderfolgen.

Ein Vergleich der Frühontogenese der menschlichen Lautgebung mit jener nah verwandter Menschenaffen läßt deutlich werden, daß hier vor allem affine Kontaklaute wirksam werden, daneben affine Distanzlaute, während die diffugen Laute zunächst noch nicht auftreten.

Die Spontanlaute des Säuglings werden gewöhnlich auch in Anwesenheit eines Erwachsenen geäußert und durch dessen Lautäußerungen gefördert, differenzieren sich aber zunächst auch bei tauben Kindern in gleicher Weise. Hier entstehen die „Urlaute“, die Lautgruppe „erre“, das „Vokalplaudern“, die „Kettenbildungen“, das „Plaudern“, die „Wiederholungen“, die „Konsonantenbildung“, die „Konsonanten-Vokal-Gruppe“ und die „Schnalzlaute.“ Die vergleichende Untersuchung legt nahe, diese Frühontogenese der Lautentwicklung des Menschen den affinen Kontaktlauten zuzuordnen, die in der Ontogenese ihre ursprüngliche Funktion hatten und hier vornehmlich in das visuelle Kommunikationssystem eingebaut wurden.

Funktionell kann bei diesen affinen Kontaktlauten auch ein autostimulatorischer Prozeß erwogen werden, bei dem die akustische Bestätigung über den eigenen Receptor auf das System in einer Form zurückwirkt, die die Tendenz zur Lautgebung fördert, wie Vergleichbares von hirnelektrischen und chemischen (Pharmaka) Auto-stimulationen ja bekannt ist. Daraus ergäbe sich eine gewisse Beziehung zum Spielverhalten, dessen Hauptfunktion auch seine Rückwirkung auf das ausübende System darstellt.

DISCUSSION

Zacharias:

Unseren Beobachtungen zufolge gibt es eine Kontakt- und eine Distanzlautgebung auch bei Säuglingen. Sie unterscheiden sich nicht nur in der zeitlichen Ausdehnung, sondern auch melodisch. Distanzlaute mit Ruffunktion zeigen eine steigendfallende Melodieform, Kontaktlaute eine fallend-steigende.

SOME ASPECTS OF SPEECH OF DEAF CHILDREN

BERNARD TH. TERVOORT

Applied phonetics and phoniatics have been interested in the speech of the deaf and the severely hard of hearing from their first decades of development. The pathology of the deaf child's speech is different from all other speech pathology insofar as the speech apparatus is normal if no multiple handicap is present, but the normal feedback control is absent due to the absence of the specific regulating and controlling auditory organ. A deaf child has to learn to speak at a later age, by kinaesthetic motor control, by visual perception of speech movements, and by accompanying vibratory or residual hearing perception. New techniques are dealing with speech teaching at a very early age, with the improved training of the combined kinaesthetic and vibratory feedback controls, and with better amplifying equipment. I shall not deal with all those here, but like to mention them by way of introduction.

I have spent the last 8 years on a diachronic study of the language development of 48 prelingually profoundly deaf children of normal intelligence and socio-economic background, 24 in the U.S.A. and 24 in the Benelux. They ranged from age 7 to 12 in 1959 and from 12 to 17 in 1965, the final year of observation. The statistics consist of 3 analyses of variance carried out on the vocabulary, 6 analyses of variance carried out on the syntax, and numerous simple computations; all of the statistics are based upon an age span of 7 to 17. The subjects were filmed on 16 mm. film with telescopic lenses, for 10 minutes each year for 6 years, in age pairs, and while engaged in private, esoteric communication. The study was sponsored by the Vocational Rehabilitation Administration of the Department of Health, Education, and Welfare, Washington, D.C., U.S.A., and by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.), The Hague, The Netherlands. Although its final report's major conclusions are of a linguistic nature, some of the observations made in its 433 pages might be of interest to phonetics too.

The total vocabulary comprises some 72,000 items. These items were not all spoken words, but were executed through: 1. speech only, 2. fingerspelling only, 3. speech with simultaneous fingerspelling, 4. speech with simultaneous signs, 5. signs only. One of the reliable measuring devices of vocabulary to be found in the literature is the ratio between the number of different words and the total number of words in a language sample, termed type token ratio (TTR): it registers the richness of a subject's vocabulary and the efficiency with which he uses it. As its applicability is limited

to samples with an equal number of tokens, and as my aim was to obtain spontaneous free samples, a new measuring device was developed mathematically, termed type token ratio value (TTRV), which could be applied to samples of unequal size.

One of the conclusions with regard to the execution of the vocabulary was, that deaf children among themselves do not use speech-only very effectively when they are young and that the ratio of this usage goes down to near zero when they grow older. On the contrary: fingerspelling is a skill that—once it is mastered—increases, while signing has a very high ratio at age 7 already. Both fingerspelling and signing, if combined with speech, raise the latter's ratio to an average to high level. Speech is not in itself a skill that cannot be mastered: it can be used effectively if it gets support from the hands and the fingers.

This is understandable if one keeps in mind that a deaf child's communication is mainly visual vs. ours which is mainly auditory: this implies that our speech is based upon oppositions serving a distinctive audibility. The deaf child's need for distinctive visibility is served much better by the movements of the hands than by the movements of the lips, but he needs to incorporate the latter because of the demands of his hearing environment and the richness of its language.

If deaf children use speech among themselves, their speech may sound different not only because of the consequences of lack of auditory feedback but more so because of the different purpose to be served. They often speak voiceless as sound is not functional for the deaf partner; they moreover speak in visual oppositions, making relevant some features which are not relevant in normal speech. The visual phoneme has been termed kineme; some phonemes which are different speech sounds, when seen are identical kinemes, termed homophenes. This goes e.g. for all phonemes executed at the same place with about the same visible movements, such as voiced vs. unvoiced consonants, or groups like $P \perp B \perp M \perp$, $K \perp G \perp X$, $T \perp L \perp R$. It is easy to see why e.g. English speaking deaf children would try to make *me-pea-bee* look different from each other by distorting the initial phonemes into different kinemes. It is also easy to see how spelling or signing clarifies the issue even better.

One final remark on the oral activity of the subjects seems to be relevant here. The voice giving or the movements of the speech apparatus could not always be termed speech: the sounds or the movements were sometimes too indistinct for that. Speech often deteriorated to some rhythmical syllabic movements of the mouth. Further research will be interested in this oral motoric activity, and might establish whether this is an intermediate stage between non-speaking and speaking in the growth toward incorporating the spoken word into the private interchange, or between speaking and non-speaking in the loss of skills, learned during school age.*

* This paper is based upon: Tervoort, Bernard T. *Final Report on Project Number RD-476-64-65 of the Vocational Rehabilitation Administration of the Department of Health, Education, and Welfare, Washington, D.C., U.S.A., titled: Analysis of Communicative Structure Patterns in Deaf Children.* etc., Groningen, The Netherlands, 1967.

ZUR PHONEMATISCHEN DIFFERENZIERUNGSFÄHIGKEIT DER VORSCHULKINDER MIT FUNKTIONELL BEDINGTEM STAMMELN

CHRISTA THEINER*

In der Literatur wird mehrfach darauf hingewiesen, daß bei Stammelern akustische Beeinträchtigungen ohne eigentliche Schädigung des Gehörs auftreten können, die sich als phonematische Differenzierungsschwäche äußern. Einige Autoren bezeichnen daher auch eine Form des funktionell bedingten Stammelns als sensorisches Stammeln.

Untersuchungen zur Ermittlung der phonematischen Differenzierungsfähigkeit der Stammer sind nur sehr wenige bekannt, meist beschränken sie sich auf das stammelnde Schulkind.

Die Erfahrungen, die in der logopädischen Arbeit mit stammelnden Vorschulkindern gesammelt wurden, führten zu der Annahme, daß bei diesen Kindern die exakte Differenzierung der Sprachlaute nicht immer gewährleistet ist.

Diese Faktoren waren Anlaß, die phonematische Differenzierungsfähigkeit der Vorschulkinder mit funktionell bedingtem Stammeln zu überprüfen.

Die experimentelle Untersuchung hatte folgende Zielstellung:

1. die phonematische Differenzierungsfähigkeit der Vorschulkinder mit funktionell bedingtem Stammeln und der normalsprechenden Vorschulkinder zu überprüfen, um festzustellen, ob der Stammer Abweichungen vom Normalkind zeigt,
2. die Häufigkeit der phonematischen Differenzierungsfehler bei den einzelnen Lautgruppen festzustellen,
3. die Beziehung zwischen Lautbildungsfehler und phonematischem Differenzierungsfehler des stammelnden Vorschulkindes zu ermitteln.

Für die Untersuchung wurde ein Prüfmaterial erarbeitet, das dem Vorschulalter gerecht wird.

Das Prüfmittel besteht aus 51 Wort-Bildpaaren. Mit diesem Material wird überprüft, inwieweit die Konsonanten der deutschen Sprache durch Vorschulkinder richtig differenziert werden.

Für die Überprüfung eines jeden Konsonanten stehen drei Wortpaare zur Verfügung. Die beiden Wörter eines Wortpaares unterscheiden sich jeweils nur in einem Phonem. Dabei ist dem zu prüfenden Laut stets der Konsonant gegenübergestellt, der als häufigster Ersatzlaut für den entsprechenden Sprachlaut bei Vorschulkindern

* Institut für Sonderschulwesen der Humboldt-Universität Berlin.

mit funktionell bedingtem Stammeln auftritt. Die Differenzierung des r-Lautes durch die Probanden wird zum Beispiel an Hand der Wortpaare Rose — Hose, Rasen — Hasen und rund — Hund überprüft.

Die Untersuchungsmethode war dem Alter der Versuchspersonen angepaßt. Dem Probanden wurde ein Wort- bzw. Bildpaar vorgelegt, nur ein Wort — das den zu prüfenden Konsonanten enthielt — wurde vom Versuchsleiter vorgesprochen. Das Kind mußte durch Zeigen auf das entsprechende Bild angeben, welches Wort es gehört hat.

Es wurden 81 Vorschulkinder mit funktionell bedingtem Stammeln und 81 normalsprechende Vorschulkinder überprüft. Das Durchschnittsalter der Probanden betrug 6 Jahre.

Das Gesamtergebnis der experimentellen Untersuchung vermittelt nachstehende Übersicht:

	Stammelnde	Normalsprechende
	Versuchspersonen	
phonemat. Differenzierungsfehler	729	175
durchschnittliche Fehlerzahl je Versuchsperson	9,00	2,16
prozentualer Anteil der falsch differenzierten Wörter je Versuchsperson	17,65	4,24

Die Vorschulkinder mit funktionell bedingtem Stammeln weisen damit im Durchschnitt das Vierfache an phonematischen Differenzierungsfehlern gegenüber den normalsprechenden Vorschulkindern auf.

Betrachtet man die phonematischen Differenzierungsfehler bei den einzelnen Versuchspersonen, zeigt sich, daß die normalsprechenden Vorschulkinder in der überwiegenden Mehrheit 0 — 4 phonematische Differenzierungsfehler aufweisen, der größte Teil der stammelnden Vorschulkinder hingegen 5—15.

Werden die phonematischen Differenzierungsfehler auf die einzelnen Lautgruppen aufgeschlüsselt, ergibt sich bei den Normalsprechenden und Stammlern folgendes Bild:

Daran ist zu sehen, daß sich bei stammelnden und normalsprechenden Vorschulkindern eine annähernd gleiche Reihenfolge ergibt. Ein Vergleich, um das Wievielfache das stammelnde Kind die einzelnen Lautgruppen gegenüber dem normalsprechenden Kind fehlerhafter differenziert, ergibt die umgekehrte Reihenfolge:

Reibelaute — Zischlaute	um das 10,1 fache
Die Laute l, R	um das 6,6 fache
Reibelaute (außer den Zischlauten)	um das 3,7 fache

Lautgruppe	Prozentualer Anteil der falsch diff. Wörter bei den Stammlern	Lautgruppe	Prozentual. Anteil der falsch diff. Wörter bei Normalsprechenden
Reibelaute-Zischlaute	12,33	Reibelaute-Zischlaute	1,22
Reibelaute (außer den Zischlauten)	14,22	Die Laute l, R	2,33
Die Laute l, R	15,5	Reibelaute (außer den Zischlauten)	3,88
Nasallaute	18,0	Nasallaute	4,88
Explosivlaute	22,55	Explosivlaute	6,22

Nasallaute um das 3,7 fache
Explosivlaute um das 3,6 fache

Aus den Fehlleistungen im Bereich der einzelnen Lautgruppen kann geschlossen werden, daß die Lautgruppen unterschiedliche Anforderungen an das phonematische Gehör des Vorschulkindes stellen.

Die Untersuchung der Beziehung zwischen Lautbildungsfehler und phonematischem Differenzierungsfehler der stammelnden Versuchspersonen ergab, daß eine Identität von Lautbildungsfehler und phonematischem Differenzierungsfehler nicht konstant gegeben ist.

Als zusammenfassendes Ergebnis der experimentellen Untersuchung kann gesagt werden, daß für Vorschulkinder mit funktionell bedingtem Stammeln eine phonematische Differenzierungsschwäche charakteristisch ist. Der Unterschied in der phonematischen Differenzierungsfähigkeit zwischen stammelnden und normalsprechenden Vorschulkindern ist statistisch signifikant.

DISCUSSION

Kress:

Es muß bezweifelt werden, daß Vorschulkinder Phoneme unterscheiden können. Das zeigen die Erfahrungen im Leselehreunterricht, wo die Segmentierung erst erlernt werden muß. Meines Erachtens handelt es sich bei der Unterscheidung von Wortpaaren wie Kanne/Tanne um Unterscheidungen von ganzheitlichen Klangbildern, solange die phonematische Struktur der Wörter noch nicht erlernt ist.

Böhme:

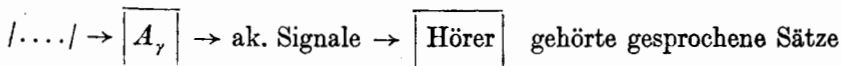
Begriff des „funktionellen Stammeln“ sollte vermieden werden. Mit Zunahme der medizinischen Erkenntnisse sinkt die Zahl der ätiologisch unklaren Dyslalien. Hinweise auf weitere Untersuchungs- und Differenzierungsmöglichkeiten; Lautagnosie—Test mit Spielzeug. Abbildungen und Strichzeichnungen sind für das Kind zu abstrakt.

KOGNITIVE VORAUSSETZUNGEN DER THEORIE EINER GENERATIVEN SIGNALPHONETIK

H. G. TILLMANN*

Auf der Tagung über generative Grammatik vor einem Jahre in Hamburg habe ich versucht, die Theorie einer generativen Signalphonetik zu entwickeln, bzw. deren Metatheorie. In den folgenden vorgeschriebenen 10 Minuten möchte ich die kognitiven Voraussetzungen einer solchen Theorie behandeln.

Unter generativen Signalphonetik verstehe ich, grob skizziert, eine Theorie, welche das Verhalten einer Signalquelle explizit macht, die solche und nur solche Signale als Zeitfunktionen erzeugt, die (über einen akustischen Wandler abgestrahlt), von einem kompetenten Hörer als gesprochene Sätze seiner Sprache identifiziert werden. Skizze (1) verdeutlicht das im Blockdiagramm:



Seien x die gehörten gesprochene Sätze und „ x “ deren symbolische Transkriptionen, so erklären wir das signalproduzierende Verhalten des Artikulators A_γ durch einen theoretisch oder praktisch konstruierbaren Erzeugungsmechanismus, der über die Symbolsequenz „ x “ operiere, die am input gewährleistet sei als hinreichende Beschreibung der phonetischen Oberflächenstruktur der zu erzeugenden Sätze. Außer den morphonematischen Symbolen, welche die Formative repräsentieren, enthalte „ x “ sowohl die satzkennzeichnenden als auch die Symbole, welche die situationsgebundene Ausprägung des Satzes kennzeichnen; die sprechercharakterischen Merkmale allerdings liefere der Artikulator selbst.

Des Erzeugungsmechanismus A_γ ist dann bereits die Theorie, die wir generative Signalphonetik nennen. Allerdings setzt diese Theorie die kognitive Instanz des Hörers voraus, der allein kompetent ist, darüber zu entscheiden, ob die Theorie richtig ist oder falsch. Darüber hinaus brauchen wir einen Hörer, der die Fähigkeit besitzt, die gehörten Sätze x als „ x “ zu transkribieren. Wenn wir den Linguisten als einen speziell geschulten Kommunikator auffassen, so können wir ihn hier einsetzen:



*IPK Bonn.

Wir benötigen dabei nur einen der beiden outputs am Linguisten, nämlich denjenigen, mit dem er sowieso seine Daten in symbolischer Form erzeugt, wobei der zweite output, der angesichts der systematischen Aspekte der Daten am 1. output die linguistische Theorie liefert, unbeanspruch bleibt. Der halbe Linguist indessen, der es als extrakommunikativer Kommunikator nur bis zum Transkribenten gebracht hat, verkörpert alle kognitiven Voraussetzungen der Theorie einer generativen Signalphonetik. Wir haben es hier mit einer heuristischen Instanz zu tun, die für die Theorie nur operational relevant ist. Es wird damit unmittelbar einsichtig, daß die kognitiven Voraussetzungen der Theorie nicht selbst Gegenstand dieser Theorie sein können.

Nachdem dieser Punkt geklärt ist, will ich zeigen, daß eine Theorie A_7 nichtsdestoweniger auch kognitiv interpretiert werden kann, indem der externe Erzeugungsmechanismus auch das kognitive Verhalten der Kommunikatoren zu erklären vermag. Dazu werde ich eine Theorie heranziehen, die von Miller, Galanter und Pribram in dem Buch „Plans and the structure of Behavior“ entwickelt wurde. Ich möchte allerdings zuvor einen kurzen phänomenologischen Exkurs über die kognitiven Tatgestände einschieben. Das entscheidende Problem, das auch in der Formulierung des Themas dieses Kongresses zum Ausdruck kommt, liegt in der Verbindung der beiden Seiten des psychophysischen Niveaus. Auf der einer Seite haben wir Signale, die sich vom externen Beobachter messen lassen, auf der anderen Seite liegt die phänomenal unmittelbar gegebene, und nur so gegebene, interne Wirklichkeit, wie sie der Kommunikator wahrnimmt und kennt. Ohne diese interne Wirklichkeit sind die externen Signale sinnlos und bloss physikalische Phänomene. Wenn wir fragen, was ein Kommunikator intern wahrnimmt, während der Wissenschaftler extern Signale registriert, so heißt die Antwort: der Sprecher nimmt seine eigenen artikulierenden Aktionen wahr, und zwar erstens taktil — propriozeptiv und zweitens auditiv, beides aber integral. Auch der Hörer nimmt nie schlechthin Schall wahr, sondern bereits kognitiv interpretierten. Der Hörende identifiziert am Schall das Verhalten einer Schallquelle, er vernimmt ihn als Türklopfen, Telephonklingen, als Geige, VW oder Trompete. Ebenso ist Sprachschall für den Hörer intern das lautwerdende artikulierende Verhalten seines Gegenübers. (Merleau — Ponty hat übrigens in seiner „Phénoménologie de la perception“ die Rolle des ‚corps humain‘ für die Wahrnehmung und im Kapitel ‚la parole‘ auch für die wahrgenommene Sprache überzeugend analysiert.) Die Frage nun, ob und wie dieses körperliche Verhalten des Sprechers, das er selbst und auch der Hörer wahrnimmt, wenn er z. B. Sätze artikuliert, bereits Phoneme, Morpheme usw. enthält, ist zunächst recht belanglos gegenüber der Tatsache, daß der Linguist das verbale Verhalten mit solchen theoretischen Mitteln beschreiben kann.

Nach diesem Exkurs wenden wir uns wieder dem Zusammenhang zwischen intern repräsentiertem Verhalten und extern meßbaren Signalen zu. Darüber hinaus muß auch der Zusammenhang zwischen dem intern repräsentierten Verhalten und dessen symbolischer Transkription „x“ geklärt werden. Wir können dem Komplex des

Artikulierens mit allem dazugehörenden neurologischen Prozessen als einen Erzeugungsmechanismus auffassen, etwa als eine Theorie A_7 . Das Produkt dieses Mechanismus, das abgestrahlte Signal, ist dann auch extern, wie Fant, Ungeheuer u. a. gezeigt haben, eine eindeutige Abbildung des extern registrierten artikulierenden Verhaltens.

Die kognitive Interpretation des Erzeugungsmechanismus A_7 muß den grundlegenden Tatbestand erklären können, daß ein kompetenter Hörer die gehörten Sätze ohne weiteres reproduzieren kann:

$$f(t) \rightarrow \boxed{H} \rightarrow f(t)$$

Nach Miller und Koautoren liegt jedem orientierten Verhalten ein kognitiver Plan zugrunde, den das Verhalten ausführt. Dieser Plan kann unabhängig vom konkreten Verhalten in Form eines symbolischen Programms beschrieben werden. Beim externen Erzeugungsmechanismus A_7 entspricht diesem Programm der steuernde Algorithmus, der über der Symbolsequenz operiert. Kognitiv interpretiert sind die Programme an die körperlichen Aktionen geknüpft, und zwar einmal als geplante, zum andern als konkret auszuführende Aktionen. Phoneme z.B. wären dann Unterprogramme in Silben- und Wortprogrammen, diese wieder Unterprogramme in syntaktischen Programmen, in welchen sie durch einen einzigen Befehl aufgerufen werden können. Der entscheidende Punkt aber liegt darin, daß die in der Programmbibliothek des Gehirns gespeicherte Information einmal als Plan und zum andern als dessen Ausführungsanweisungen aufgefaßt werden können. Das hängt übrigens davon ab, auf welcher Ebene der Unterprogramme der kognitive Plan gerade repräsentiert wird. Der Körper kann eben die eigenen Aktionen wahrnehmen, sich daran erinnern, sie antizipieren, d. h. planen, und sie, dem Plan folgend, wieder ausführen.

Um das Verhalten des Hörers auf dem Boden dieser nur knapp skizzierten Theorie zu erklären, müssen wir den Hörer als einen Mechanismus verstehen, der anhand der einlaufenden Signale, die ja das Abbild des artikulierenden Verhaltens sind, dieses Verhalten synchronisiert, derart, daß er zugleich die zugrundeliegenden Programme erzeugt. Der Linguist schließlich hat für seinen 1. output eine Art Compiler entwickelt, der es ihm gestattet, die internen Programme in andere interne Programme zu übersetzen, denen die externen graphischen Symbole so zugeordnet sind, daß man die Programme sogar lesen kann. Ich hoffe, daß mit diesen wenigen Andeutungen das Gerippe einer möglichen Kognitiven Theorie des verbalen Verhaltens deutlich geworden ist.

ОБ ЭКСПРЕССИИ ЗВУКОВ

В. ТОМАНОВИЧ

Хотя необычайность звуковой формы слова и привлекает внимание, а звонкость и гармония звуков создает приятную импрессию у слушающего, всё же представление об их природе и функции было бы неполным и неверным, если бы при исследовании экспрессивность смешивалась с необычайностью и пустой гармонией звуков, без какой-либо связи с содержанием.

Факт, что в разных языках в известной мере употребляются различные элементы фонологической экспрессии, а подчас один и тот же звук придает слову в одном случае одну, а в другом другую эмоциональную окраску, например, смягченный звук в одном случае гипокористическую, а в другом пейоративную, мог бы при поверхностном наблюдении повлиять на понимание связи между значением и звуком как произвольной и случайной и привести таким образом к недооценке её.

Такому взгляду способствовало бы и неучитывание различия фонетических систем отдельных языков и их истории, как и выпущение из виду факта, что музыкальное выражение в слове выступает вместе с его интеллектуальным значением, выражая, как отражение мимического выражения, лишь одну из особенностей высказываемого объекта мышления или один из компонентов мимического выражения чувства (в упомянутом случае ослабление мышц посредством перенесенной на голос ослабленной артикуляции), из чего следует, что исполняет функцию, подобную функции качественного имени прилагательного или наречия; вследствие чего может относиться к различным понятиям, а потому находится и в словах, вызывающих различные чувства. Мягкое мимическое выражение может, как и прилагательное „мягкое“, обозначать свойство нежных объектов, нежные чувства, но также и свойство гнилых, гнойных материй в разложении, затем дряблых, худосочных живых существ и т. п., как и мимическое выражение чувства брезгливости, отвращения.

Ударение слова становится экспрессивным только постольку, поскольку ради экспрессии отклоняется от шаблона. Среди следов старой, прежней экспрессии кроме вокатива *páter* < *patēr* и др., W. Wundt, *Die Sprache*, можно привести и сербскохорватский пример *дче* — *от̄ац* (лит. *д̄тац*). Наряду с примером латинского императива *dic, duc, fac*, O. Jespersen,

Die Sprache, без конечного -е, императив в славянских языках без конечного -и, нир. поль. *nieś*, сербскохорв. екср. *иде, реч, носѝ*, лит. *пѝси* и др. Наряду с греческим *τροπέω, τροπάω*, слав. *nesq, nositi, naša.ti* < **nōsjati*. Год. сборник фил. фак., Скопье, 1960 г.

Интонация вопросительного междометия *э* или *а* одинакова в главной характеристике с ударением вопросительного предложения, таким образом и из этого видна их общая принадлежность к фонетической экспрессии. И когда в вопросительном предложении находится вопросительное местоимение, и тогда оно произносится с вопросительной интонацией. Мимическое выражение ожидания при вопросе настолько сильно, что его отражение на высказываемом не может отсутствовать. В старогреческом языке акут вопросительных местоимений *τίσ?* кто? *τί?* что? и тогда, когда находится в позиции, в которой должен был бы измениться в гравис, оставался, как единственное отступление от правила, всегда неизменным. Тут акцентная интонация не подчинилась фонетическому правилу, вследствие своей экспрессивной функции.

Вследствие их общего свойства давать непосредственное выражение, фразовым музыкальным выражением охвачены и составляют с ним одно целое все музыкальные элементы (тон, тембр, динамика, ритм и темп), невзирая на то, выражают ли они чувства, волю или же исполняют какую-либо логическую функцию, и лишь тогда, когда они рассматриваются таким образом — целостно, они нам представляются в своей действительной и полной функциональной ценности.

RELEVANZ DER GESTALTELEMENTE DER SLOVENISCHEN TONEME

JOŽE TOPORIŠIČ

Besonders den Slavisten ist es wohlbekannt, daß die slovenische Sprache (das ist die am weitesten im Westen gelegene südslavische Sprache) über einen Tonem-gegensatz verfügt: bei demselben Lautbestand kann man nämlich nur auf Grund der verschiedenen Akzentgestaltung Worte voneinander unterscheiden; vgl. z. B. *dán* (= gegeben) ≠ *dán* (= der Tag), *ꝑót* (der Weg) ≠ *pót* (der Schweiß) oder *brátov* (Bruders, vom Bruder) ≠ *brátov* (= der Brüder). Den Akzent /' / nennt man Akut, den /˘ / Zirkumflex.

Sieht man sich die bisherigen Aussagen¹ der Phonetiker und Sprachwissenschaftler über die Natur dieser Akzente näher an, stellt man gleich fest, daß es da 3 Elemente gibt, die als phonologisch relevant betrachtet werden können:

1. die kynetische Ausgeprägtheit der 1. oder der 2. Hälfte des Tonem-Trägers;
2. der Tonverlauf des Tonemträgers und
3. die relative Realisationshöhe des Tonemträgers.

Graphisch dargestellt:

1. — = ≠ — —

2. / ≠ \

3. ... ≠ ...

Die erste Ansicht vertrat die älteste slovenische Tradition (Stanislaus Škrabec), experimentell bestätigt von Olaf Broch. Die zweite Auffassung vertrat, nach Mazing, Matija Valjavec und nach ihm noch viele andere (z. B. auch Olaf Broch); experimentell wurde diese Auffassung bewiesen von Fran Bezljaj. Die dritte Auffassung kommt von Olaf Broch; der Tonhöhenunterschied war ihm relevant in gewissen Satzstellungen und beim schnelleren Sprechtempo. Den Tonhöhenunterschied erklärte zum alleinigen Unterscheidungsmerkmal Mitja Sovre, was später von Bože Vodušek auch experimentell bestätigt wurde.

Welcher Ansicht kann man glauben schenken?

¹ Jože Toporišič, Pojmovanje tonemičnosti slovenskega jezika; *Slavistična revija* (Ljubljana) XVI, 1967, 64—108.

Es kommt noch hinzu, daß es nicht klar ist, ob die Ausgeprägtheit der Teile des Tonemträgers durch die Intensität oder durch die Tonhöhe erlangt wird. (Škrabec dachte an den Intensitätsunterschied, Broch behauptet, daß die Intensität mit der Tonbewegung parallel ist, Trubeckoj war nur verschiedene Behandlung der Teile des Tonemträgers wichtig). Genau so ist es noch unentschieden, ob die Tonemopposition an der betonten Silbe allein oder an der folgenden Silbe (Silben) realisiert wird, oder an beiden zugleich. Und schließlich geht es bei der Auffassung vom Tonhöhen Gegensatz darum, ob die Relevanz der Realisationshöhe der verschiedenen betonten Silben-träger zukommt oder aber dem Verhältnis der betonten gegenüber der unbetonten nachfolgenden Silbe.

Wie aus dem Titel meines Berichtes hervorgeht, glaubte ich, daß es sich bei Tonengensatz um ein Gestaltproblem handelt, d.h. daß mehrere Elemente der Tonemgestalt relevant werden können. Nämlich: 1. die Tonhöherealisation der Arsen und Thesen akutiert bzw. zirkumflexiert betonter Worte, 2. das Tonhöherealisationsverhältnis zwischen der Arsis und Thesis im Rahmen eines auf eine bestimmte Weise tonierten Wortes und 3. der Tonverlauf selbst (wie schon von Broch behauptet wurde). Jetzt bin ich zur Überzeugung gelangt, daß dem Tonhöhenunterschied allein die Relevanz zukommt. Zu dieser Ansicht bin ich gekommen, nachdem ich Messungen an verschiedenen Stellen des Grundtonverlaufs der Toneme vorgenommen habe: in der Mitte des Tonverlaufs der Arsis (das ist die Silbe des Wortes, an der in den Wörterbüchern der Akzent verzeichnet wird) und Thesis (das ist die der Arse nachfolgende unbetonte Silbe desselben Taktes). Ebenso habe ich auf die Form des Tonverlaufs der Arsis und ihren Inklinationwinkel Acht gegeben, zuletzt auf die Realisationstonhöhe der nacheinander folgenden Arsen und Thesen.

Der Unterschied zwischen dem Akut und Zirkumflex wird demnach durch folgende Elemente aufrechterhalten:

Der Akut ist durch eine gegenüber dem Zirkumflex niedriger realisierte Arsis und höher realisierte Thesis gekennzeichnet, also etwa /_· ≠ —./ . Bei den oxytonierten Worten vor einer Pause übernimmt die Rolle der Thesis der zweite Teil der Arsis. Da aber der Tonverlauf natürlich ununterbrochen ist, kann die Tonhöhenopposition nur durch einen (nach Satzpositionen verschiedenen) Tonverlauf der Arsis erreicht werden, was die Vermutung aufkommen läßt, daß in solchen Fällen die Form des Tonverlaufes selbst relevant sein könnte. Demnach hätten wir bei den baritonierten Worten (und bei den oxytonierten mit der nachfolgenden vorbetonten Silbe des nachfolgenden Wortes als Thesis) mit dem Tonhöhenunterschied zu tun, bei den oxytonierten jedoch mit der Tonverlaufopposition. Dies trifft jedoch nicht zu, denn auch bei den Oxytonen vor einer Pause ist nicht der Tonverlauf selbst relevant, sondern eben verschiedene Realisationsebenen der beiden Teile der betonten Arse.

Wird der Unterschied in der Realisationsebene (-höhe) der Arsen neutralisiert, so bleibt als relevant der Unterschied in der Realisationshöhe der beiden nachfolgenden Thesen. Beim Akut ist sie höher realisiert als beim Zirkumflex.

Meßwerte der Unterscheidungselemente sind nicht etwa konstant, sondern stark

veränderlich. Der Tonhöhenunterschied ist am größten im ersten Takt eines Satzsegmentes (und bei besonderer Hervorhebung), vermindert sich aber gegen das Ende des Segmentes immer mehr, und in gewissen Stellungen (das sind satzphonetisch wichtige Punkte: die Kadenz, Antikadenz, und Halbkadenz) kommt zuweilen zur sehr großen Verminderung des Gegensatzes oder sogar zur phonetischen Neutralisation zwischen Akut und Zirkumflex. Phonologisch bleibt das tonematische Merkmal der Worte jedoch erhalten, was durch eine andere Sprachumgebung bewiesen werden kann.

Alles das erschwert das Identifizieren der slovenischen Toneme. Will man das tonemische Sprechen erlernen, so muß man alle diese verschiedenen Allotone gut beherrschen lernen. Hat man nur eine typische satzphonetisch bedingte Tonemform gelernt, kann man die Toneme in den anderen Positionen kaum oder gar nicht identifizieren. Ich glaube, daß man sich typische Identifizierungsparameter für einzelne Takte der Satzsegmente erwerben muß, sonst findet man sich im Sprechstrom gar nicht zurecht.

Zum Schluß möchte ich noch zusetzen, daß diese Identifizierungsparameter für verschiedene Dialekte etwas verschieden sein können, auf Grund wessen dann auch zu Fehlidentifikationen der Toneme kommt.

Beweisen kann man alles von mir Gesagte natürlich nur in längeren Aufsätzen. Sie warten auf die Veröffentlichung in einer slovenischen slavistischen Zeitschrift.²

² Jože Toporišič, Liki slovenskih tonemov; *Slavistična revija* (Ljubljana) XVI, 1968, 315—393.

АКУСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ СМЫСЛОВОГО ЧЛЕНЕНИЯ ПРЕДЛОЖЕНИЯ

И. Г. ТОРСУЕВА*

Наблюдения над речью позволяют выделить, наряду с актуальным членением предложения, еще один вид смыслового членения. Основой этого членения является неравномерное распределение смысловой важности в предложении. Актуальное членение базируется, в основном, на интеллектуальном моменте, т. е. на ходе мысли от известного к неизвестному. Понятие смысловой важности включает в себя одновременно эмоциональный момент, что очень существенно для бытовой разговорной речи. Кроме того, многие предложения с точки зрения актуального членения рассматриваются как нерасчлененные высказывания („Byl jednou jeden král“, пример В. Матезиуса), тогда как с точки зрения смысловой важности их части неравноценны. Схематически это предложение можно изобразить как двух-

jeden král,

byl jednou] _____ где вторая часть имеет большую степень смысловой важности, чем первая. Весьма часто ядро высказывания сопровождается сопутствующими словами, которые менее важны для говорящего и слушающего, чем собственно ядро, однако в рамках актуального членения этот момент не учитывается. Помимо этого, смысловая важность как основа членения позволяет делить предложение на любое число частей.

Таким образом, членение по смысловой важности дополняет актуальное членение предложения.

Элементами выделяемого нами членения являются отрезки предложения, наделенные различной степенью смысловой важности. Последовательность элементов может быть различной. Наиболее важный элемент может находиться рядом с наименее важным, со вторым по степени важности и т. д.

На графике представлены возможные трехэлементные шкалы.



(Больший подъем на графике, соответствует большей степени важности.)

* ЛЭФИПР 1 МГПИИЯ, Москва.

Границы смысловых элементов могут совпадать или не совпадать с границами различных единиц, а именно — слова, члена предложения, ритмической группы и т. д. Смысловый элемент может охватывать лишь часть члена предложения или включать в себя несколько членов предложения.

Элементом смыслового членения в речи соответствуют постоянные характеристики в системе каждого языка, в частности — в системе интонационных средств. С целью выявления этих характеристик был проведен эксперимент, который заключался в записи дикторами — носителями языка примерно 500 фраз на французском, испанском и итальянском языках и последующей расшифровке интонограмм. Фразы содержали не более четырех смысловых элементов. Анализ экспериментального материала показал, что наиболее ярко выраженные показатели имеет наиболее важный элемент (V_1) а также второй по степени важности элемент (V_2).

V_1 характеризуется

во французском языке — частотным пиком, наиболее широким диапазоном по частоте основного тона, наибольшим интервалом падения или повышения частоты основного тона;

в итальянском языке — частотным пиком (или наиболее высоким частотным уровнем), наиболее широким диапазоном по частоте основного тона, наибольшей амплитудой интенсивности;

в испанском языке — частотным пиком, наибольшим интервалом падения или повышения частоты основного тона, а также увеличением средней длительности слога. V_2 характеризуется

во французском и испанском яз. — наибольшей амплитудой интенсивности;

в итальянском языке — увеличением средней длительности слога.

Эти показатели до известной степени варьируют в каждом из перечисленных языков как по смысловым вершинам, так и по образуемым им смысловым шкалам.

При сравнении структурных показателей V_1 во французском, испанском и итальянском языках следует отметить, что частотный пик характеризует наибольшую степень важности во всех трех языках. Французский язык по структурным показателям занимает как бы промежуточное положение между испанским и итальянским.

Полученные структурные характеристики составляют модели смыслового членения по степени важности. Эти характеристики не совпадают с характеристиками членов предложения, которые не имеют собственных интонационных показателей. Этот факт свидетельствует о том, что элементы смыслового членения являются единицами особого уровня системы языка, который условно можно назвать супрасинтаксическим. К этому же заключению подводит признание особой основы смыслового членения, а также несовпадение границ смыслового членения с границами членов предложения и других единиц.

CUES TO THE RECOGNITION OF SOME LINGUISTIC FEATURES OF WHISPERED SPEECH IN ENGLISH

J. TRIM

The study of whisper, as of other abnormal speech mechanisms, is of interest for the understanding of the relation of speech to language. If it is considered that speech perception is a matter of an automatic response to certain acoustic constants, whether absolute or relational, in the speech wave, necessarily produced by automatized articulations, phonological units will be defined in strictly phonetic terms. In the alternative view, it may be held that, where speaker and listener share the same linguistic competence, it is the task of the speaker under abnormal conditions to incorporate into the speech-wave he produces in a particular performance sufficient information for the listener to identify the text intended; and it is the listener's task to make that identification on the basis of the clues available.

In whisper, vocal cord vibration is avoided, the periodic voice source being replaced by an aperiodic voice source. The energy supplied over a wide frequency band is filtered in accordance with the resonant characteristics of the vocal tract, as in voice. The spectral features characteristic of many speech-sounds are thus unchanged. The first and second vowel formants, for instance, are as well-marked and perhaps more stable for not being subject to fluctuations in harmonic intervals and the same applies to consonants primarily characterized by formants and formant transitions. The local supraglottal noise effects characterizing plosives; affricates and fricatives are unaffected, but the question of the voiced-voiceless distinction arises. In most cases, concomitant features, well-known from phonetic descriptions, take over.

In final position the cues are: length of preceding vowel, lateral or nasal; in the case of plosives: length of stop, length and intensity of the noise burst or release; in the case of fricatives; the length and intensity of friction. Intervocally in disyllables the length of the stop or friction is dominant. If no substantial degree of lengthening is present an intended "voiceless" consonant is likely to be heard as "voiced" e.g. "latter" as "ladder". Initially the cues are again length and intensity, and aspiration. In whisper aspiration cannot of course be equated with delayed onset of voice. Instead, the onset of a vowel-like segment with strong concentration of energy in the formant is separated from the initial consonant by a segment with diffuse noise centred on 4 kc. The presence of this diffuse noise, as opposed to the abrupt onset of whisper with clear formant structure (especially initiated by a glottal plosive) characterizes h-words as opposed to V-words. There do not appear to be cases in

which the use of whisper for voice leads to lexical homophony. The residual features employed are often accentuated in whispered speech, but still the demands on the listener are increased, with a resultant lowering of recognition scores.

By comparison with such lexical features, the intonation of whispered speech presents particularly interesting problems. It has been asserted that the concept of intonation is inapplicable to whisper, since the vocal cords do not vibrate, and the speech wave, being aperiodic, has no fundamental frequency. It is said that in that case the voice has no pitch, and that tone and intonation are excluded. Returning, however, to the first point made in this paper, we may reply that if intonation is part of the meaningful linguistic form of the utterance, speaker and listener will readily adapt to an alternative manifestation.

To test this, a small experiment was conducted along the following lines. Two lists of sixty words each were constructed, in which each member of one of two sets of ten monosyllables, to some extent phonetically balanced, was combined with each of 6 nuclear tones (low fall, high fall, low rise, high rise, fall-rise and rise-fall) in an arbitrary order. The lists were recorded by the experimenter in the order: set I, voiced (test A); set II, whispered (test B); set II, voiced (test C); set I, whispered (test D). The sequence of tones was arbitrarily different in each list. A small panel of trained, professional phoneticians were then asked to write down immediately against the printed word the tone they thought to have heard. The words were presented on tape, at a rate which would allow just enough time for the subject to note the tone, but not to reflect. One or two subjects complained of the speed and may conceivably have gained higher scores if allowed longer. After a short rest, the subjects repeated the test, to see whether any short-term learning had occurred. Two of the subjects then repeated the whole test after six weeks, to see whether the learning effect, if any, had persisted.

The results of the experiment are summarized in the following tables:

I. Tones correctly recognized:

Subject	A1	B1	C1	D1	A2	B2	C2	D2	Total (480)
A	45	30	49	22	52	29	49	29	305
B	49	33	51	27	53	29	53	33	328
C	50	26	52	24	52	36	58	34	332
D1	53	31	57	30	58	40	57	30	356
E	58	32	59	31	58	44	60	38	380
F1	58	39	60	38	57	46	60	45	403
G	59	46	60	45	59	49	59	54	431
D2	57	36	58	39	57	44	60	40	391
F2	56	50	60	44	56	50	60	44	420
Total	485	323	506	300	502	367	516	347	

II. Breakdown of recognition scores by word and tone:

	no	there	now	oh	yes	six	would	fine	loose	out	Total (180)
\1	5	11	17	18	16	11	12	11	15	10	126
\2	14	17	15	13	15	15	4	16	14	13	136
/3	15	6	14	12	18	9	18	13	4	6	115
/4	3	3	4	2	10	8	6	1	11	1	49
√5	12	15	14	15	14	17	13	9	12	13	134
^6	17	13	12	9	14	15	14	9	16	11	130
Total (108)	66	65	76	69	87	75	67	59	72	54	690 (774)

b) Test d

	so	nine	you	nice	will	Franks	oh	where	shut	good	Total
1	17	12	18	11	15	14	14	16	13	7	137
2	7	14	6	13	7	17	14	8	16	7	109
3	1	10	15	17	11	5	14	18	12	13	116
4	5	5	4	4	5	2	4	13	2	4	48
5	14	8	13	8	17	6	17	13	13	14	123
6	15	16	14	13	8	9	14	6	9	12	116
Total	59	65	70	66	63	53	77	74	65	57	649 (752)

c) Scores for attempt 1: 1614 (max 2160)
attempt 2: 1732

d) Scores for test A. 987 (max 1080)
B. 690
C. 1022
D. 647

e) Scores for Voice 1 991 (max 1080)
Whisper 1 623
Voice 2 1018
Whisper 2 714
Voice 1 + 2 2009 93% (max 2160)
Whisper 1 + 2 1337 62% of max, 66% of voiced
Whisper 1 + 2 given interchangeability of high and low fall 1526 75%.

f) Long Term effect: subjects D and F, take 1 759 (max 960)
take 2 811

Though the number of subjects involved in the experiment is small, certain trends seem clear and relatively constant. Small but constant differences of perfor-

mance between subjects persist throughout the test, reflecting the relative skill and experience of the subjects in this field. A small but steady improvement of performance occurs in the voiced tests, flattening off as the 100 % level is approached or reached. The recognition scores for whisper are lower, but far too high, in all cases, to result from chance. The improvement between the first and second attempts is steeper for whisper. There was, however, a slight drop between tests B and D, on both test and repeat. A fatigue effect is perhaps responsible. The figures for the retake, 6 weeks later, by subjects D and F show that the learning effect had persisted.

It should be emphasized that the subjects received no feedback concerning their performance, so that learning proceeded only by the subject sensitizing himself to the recognition cues present.

Some error patterns in whisper recognition were very consistent. The nature of the lexical material had a restricted effect, more attributable to the frequency of use of words in monosyllabic utterances than to their phonetic consistency. Low and high falls were rarely falsely identified except as each other. Indeed, if in view of the gradient nature of falling intonations we allowed the interchange of whispered falls, their recognition rate would rise to 96 % for test B and 97 % for test D. A particularly striking feature of the results is that the high rise, with recognition scores of 96 % in test A and 97 % in test B, has a score in whisper of only just over 25 %, being mainly identified as high fall, low rise or rise fall.

Spectrograms were then made of the 240 items on the test, and examined in an attempt a) to identify the cues to the recognition of tones, b) to relate the distribution of judgements for particular items to the variation in the acoustic properties of the utterance, c) to determine whether the cues were ones present in voiced speech, or new features specific to whisper.

There is no space to deploy the evidence in detail here. The apparent answers to the questions are that Meyer-Eppler's observations, i.e. that the amplitude contour has the shape of the perceived pitch curve, and that the frequency of the third formant rises and falls in conformity with perceived pitch are largely confirmed, though 3rd formant bending is often minimal, and disregarded by listeners when it is in conflict with amplitude contour. This latter feature exists to some extent in voiced speech, as a modification of a general steady diminution of energy in the vowel segment. In whisper, the features are heightened, though still not entirely cancelling out the diminution. It is this feature listeners appear to rely on initially, which accounts for the much higher recognition rate of falls than rises, and the identification of the high rise as high fall or rise fall.

In addition to the features noted, use is made of further spectral features. In some cases, the vowel formants are considerably affected, as in "you". More often, the region up to 2.5 kc is relatively unaffected, and pitch effects are produced by fluctuation in the amount of energy contained in formants 3, 4 and 5, between 2.5 and 4—occasionally—5 kc. The presence of amounts of energy in that area, usually concentrated in formant bands, but sometimes more diffuse, occurs in high pitched

whisper and its absence in low pitched whisper. Falls, rises, fall rises and rise falls are cued by successions of these states. Low fall and high fall, low rise and high rise are differentiated by the relative intensity of this high frequency noise. It is to these features, absent from the spectrograms of the voiced items in the tests, that listeners are learning to respond, rapidly and, it seems, with persistent effect, as their recognition scores improve.

BERICHT ÜBER DIE ERGEBNISSE DER ENQUÊTE ZU JAKOBSON'S GESETZ VOM SCHICHTENBAU DES SPRACHLAUTSYSTEMS

FELIX TROJAN

Auf dem Kongreß in Münster (1964) wurde von der „Internationalen Vereinigung für Biophonetik“ die Veranstaltung einer Enquête zu Jakobsons Gesetz vom Schichtenbau des Sprachlautsystems beschlossen. In den an rund 400 Fachleute in allen Teilen der Welt ausgesandten Unterlagen wurden statistische Angaben über die Fehlbildung von Lauten bei normalhörigen Kindern und bei Aphasischen erbeten. Bisher wurden 25 Aufstellungen eingesandt und eine Reihe weiterer in Aussicht gestellt. Dieses Material bezieht sich fast ausschließlich auf Dyslalien, so daß eine weitere Enquête über aphasische Laustörungen in Angriff genommen werden mußte. Die Enquête über die Dyslalien soll fortgesetzt werden.

Die Dyslalien stellen qualitative Merkmale dar, die einer bestimmten Gruppe der Bevölkerung — eben den Dyslalikern — zukommen. Quantitativ sind gegeben: die Zahl der untersuchten Kinder, ihr Alter und die Anzahl der festgestellten Dyslalien. Diese Werte sind jedoch in den eingelangten Aufstellungen (zum großen Teil gegen die ausgesandten Anweisungen) nach verschiedenen Gesichtspunkten ausgewählt. Was die Zahl der Kinder betrifft, wurden in einem Fall nicht nur die Dyslaliker, sondern auch die sprachgesunden Kinder einer schulischen Organisation einbezogen. Dadurch kam auf das einzelne Kind ein Mittelwert von nur 0,07 Dyslalien. Es kann aber auch die Altersspanne variieren. So wurden einmal nur Kinder von 2—3a, ein anderes Mal solche von 6—14 Jahren untersucht. (Die ausgesandte Anweisung hatte in dieser Hinsicht keine Richtlinien gegeben). Da die Stammelfehler zunächst physiologisch sind und erst von einem bestimmten Alter an pathologisch werden, nehmen sie mit zunehmendem Alter ab (Shainemann, 1938).

Was nun die Zählung und die Art der Dyslalien betrifft, so wurden in dieser Hinsicht sehr verschiedene Gesichtspunkte eingenommen:

1. wurden in eine Stichprobe entweder alle Arten von Dyslalikern aufgenommen oder nur solche mit multiplen oder nur mit partiellen Dyslalien. Natürlich erhöht sich im ersteren Fall der Mittelwert.

2. wurden von den meisten Beobachtern nur diejenigen Laute in die Statistik aufgenommen, die durchgehends fehlen, fehlgebildet oder substituiert wurden. Einige wenige Beobachter aber haben alle Auslassungen, Fehlbildungen und Substitutionen gezählt, die sich im An-, In- und Auslaut finden, wodurch sich wieder der Mittelwert erhöht.

3. spielt auch das Maß der sozialen Betreuung eine Rolle. Eine Untersuchung an 300 Kindern in Linz, Oberösterreich, hat ergeben, daß die Abweichung in der Zahl der Dyslalien bei sozial gut betreuten Kindern gegenüber schlecht betreuten nach dem Chi²-Test stark signifikant ist.

4. wird die Zahl der erfaßten Dyslalien dadurch beeinflußt, daß in manchen Tabellen neben den Auslassungen und Substitutionen die Fehlbildungen nicht oder nur in geringem Maß berücksichtigt sind. Und

5. ist die Zahl der Dyslalien auch ganz besonders abhängig von der Sorgfalt und der Strenge der Untersuchung. Genau genommen, ließen sich Vergleiche zwischen Stichproben nur bei gleichen Untersuchern anstellen, was bei internationalen Untersuchungen allerdings praktisch unmöglich ist.

Über die bei den einzelnen Stichproben angegebene Anzahl von Kindern, das Lebensalter und die Zahl der Dyslalien hinaus läßt sich auch noch eine freilich nicht exakt bestimmbare zeitliche Reihenfolge der einzelnen Phoneme im Sinne der Theorie von R. Jakobson und seiner Vorgänger sowie der amerikanischen Developmental Charts (wie sie Poole-Blanchard und Wellman-Case aufgestellt haben) annehmen, die freilich schon bei den einzelnen Kindern, in bestimmten Punkten auch bei den einzelnen Beobachtern, verschieden ist. So stellen z.B. die amerikanischen Beurteiler die palato-velaren Plosive vor die labiodentalen Frikative, während die europäischen die umgekehrte Reihenfolge bevorzugen. Bei der Klärung solcher Fragen müßte vor allem festgestellt werden, ob der Beobachter die genetische Reihenfolge, die phonologische Ordnung oder den Schwierigkeitsgrad im Auge hat.

Eine Zusammenstellung der Anzahl der Dyslalien bei den Aufstellungen einzelner Phoneme auf Grund von 14 Aufstellungen und einer Gesamtanzahl von 2433 dyslalischen Kindern zeigt einen steilen Anstieg der Werte von den frühen zu den späten Lauten. Wo schon bei den frühen Lauten (*M, P, B, N, T, D*) etwas erhöhte Werte auftreten, hat dies seine Gründe darin, daß a) die plosiven Lenes „noch“ als plosive Fortes auftreten ($b = p, d = t$), b) eine interdental Bildung besonders der Dentalen *N, T, D* miteingerechnet wurde (Interdentalität ist bekanntlich ein typisches Symptom frühkindlicher Sprachentwicklung) und c) bei Kindern mit einer slawischen Muttersprache Palatalisierungsfehler einbezogen wurden. In der Tabelle I (und nur in dieser) wurden die Werte hinzugefügt, die sich ohne diese Störungen ergeben würden. Was im besonderen den Punkt a) betrifft, stellt die stimmhafte Bildung der Plosiven eine gewisse Schwierigkeit dar, die Annahme liegt nahe, daß es sich um entwicklungs geschichtlich jüngere Laute handelt, eine Ansicht, die H. Abrahams schon 1954 geäußert hat. Tabelle II zeigt im allgemeinen mittlere, Tabelle III vorwiegend hohe Werte.

Die biophonetische Deutung dieser Tatsachen geht von der 1952 von mir experimentell gewonnenen Feststellung aus, daß die am frühesten auftretenden Laute gerade diejenigen sind, die auch beim stimmhaften Kauen hervorgebracht werden. Es sind die Plosive und Nasale des ersten und des zweiten Artikulationsgebietes (wobei die Plosive vielfach noch in der urtümlichen Form der Schnalze erscheinen, was mit

Tabelle I

Dyslalische Kinder	Nation	Alter	Zahl	<i>m</i>	<i>p</i>	<i>b</i>	<i>n/n</i>	<i>t/t</i>	<i>d/d</i>
tschech.	2—3a	100	—	3	13	7/13	10/11	14/15	(a: 5, c: 0)
slowak.	3—6a	100	—	—	11	8/8	13/7	12/12	(a: 8, b: 3, c: 2)
deutsch	3—11a	59	—	3	10	6	10	7	
poln.	3—12a	100	3	5	25	5	2	29	(a: 2)
franz. (belg.)	3,6—8a	115	—	3	—	5	7	6	(a: 3)
deutsch	4—5a	100	2	—	1	1	1	—	
franz.	4—14a	101	2	3	19	5	6	13	(a: 4)
tschech.	4—16a	100	—	—	2	1/6	13/6	6/7	(a: 3)
deutsch	5—6a	200	1	1	1	—	1	1	
deutsch	5—6a	216	3	5	2	11	8	17	
deutsch	5—6	300	—	—	—	4	11	7	
deutsch	6—14a	942	1	2	2	7	14	14	(b!)
zusammen:				2433	dyslal. Kinder	Angabe in Prozenten			

Anm. Die Angaben in Klammern geben die Zahl der Dyslalien an, die sich bei Nichtberücksichtigung der folgenden Fehler ergeben:

a: ohne $b = p, d = t$; b: ohne interdental Bildungen; c: ohne Palatalisierungsfehler.
Wo neben den nichtpalatalisierten Lauten auch palatalisierte angegeben sind (z.B. d/d,) beziehen sich die indices nur auf die nichtpalatalisierten Laute.

der Anschauung von R. Stopa übereinstimmt). Die Tatsache, daß die Plosive des dritten Artikulationsgebietes den Kindern vielfach große Schwierigkeiten bereiten, läßt sich nicht phonologisch, sondern nur in diesem Sinne erklären. Ebendiese Verankerung in der Mastikation (als der Primärfunktion gegenüber der Sekundärfunktion der Artikulation) läßt daran denken, daß — in Übereinstimmung mit peripheren Unterschieden — die alten Laute in höherem Maß in der extrapyramidalen Motorik wurzeln, während die späten und besonders die schweren Laute mit ihrer Feinmotorik in weit höherem Grad in der kortiko-bulbären Motorik verankert sind. Dieses zuerst von dem Anatomen H. Hayek geäußerte Aperçu erlaubt es, den rein phonologischen Erklärungsprinzip, das sich durch die Strukturverschiedenheiten der Sprachen notwendig aufspalten muß und darum nicht zum Träger einer für alle Sprachen gültigen Einheitlichkeit werden kann, ein biologisches zugrundezulegen, das die Gültigkeit der phonologischen Feststellungen nicht aufhebt, wohl aber dessen

Tabelle II

Dyslalische Kinder			<i>l/l.</i>	<i>j</i>	<i>h</i>	<i>f</i>	<i>v</i>	<i>k</i>	<i>g</i>	<i>ç, x</i>
Nation	Alter	Zahl								
tschech.	2—3a	100	53	10	31	29	47	23	28	29
slowak.	3—6a	100	14/1	1	8	4	3	7	12	3
deutsch	3—11a	59	3	16	2	22	14	51	41	—
poln.	3—12a	100	5	4	29	14	16	35	67	—
franz.	3,6—8a	115	3	—	—	14	9	16	20	—
(belg.)										
deutsch	4—5a	100	7	1	—	7	2	24	12	12
franz.	4—14a	101	3	—	—	29	26	14	27	—
tschech.	4—16a	100	11	—	1	3	1	6	1	1
deutsch	5—6a	200	1	3	—	5	3	16	7	12
deutsch	5—6a	216	15	19	—	16	17	28	32	83
deutsch	5—6a	300	1	—	2	4	—	9	10	4
deutsch	6—14a	942	7	4	1	1	1	8	8	7
zusammen:			2433 dyslalische Kinder		Angabe in Prozenten					

Tabelle III

Dyslalische Kinder			<i>r</i>	$\frac{rf}{r\bar{z}}$	<i>s</i>	<i>z</i>	<i>ts</i>	<i>s,</i>	<i>ts,</i>	<i>z,</i>	<i>tf</i>	<i>f</i>	<i>z</i>
Nation	Alter	Zahl											
tschech.	2—3a	100	100	100	64	67	69				68	29	38
slowak.	3—6a	100	57		48	49	34				45	56	55
deutsch	3—11a	59	31		49	—	20				—	—	—
poln.	3—12a	100	72		46	30	49	17	20	12	54	80	73
franz.	3,6—8a	115	14		54	43	—				—	77	57
deutsch	4—5a	100	44		61	—	—				—	54	—
franz.	4—14a	101	18		34	41	—				—	64	74
tschech.	4—16a	100	60	83	58	55	61			48	45	47	—
deutsch	5—6a	200	21		38	—	—				—	36	—
deutsch	5—6a	216	25		62	—	—				—	72	—
deutsch	5—6a	300	17		52	—	50				—	62	—
deutsch	6—14a	942	29		72	—	—				—	44	—

allgemeingültige, in die Lebenstiefen hinabreichenden Voraussetzungen klar herausstellt.

Abschließend sei noch allen Mitarbeitern an der Enquête der herzlichste Dank ausgesprochen. Es waren dies: Pater K. G. Audenaert, Kermt, Belgien; Prof. Dr. R. Becker, Berlin; Prof. Dr. P. Biesalski, Mainz; Dr. A. Chmelová, Praha; Dr. J. Dvončová, Bratislava; E. Ellis und R. R. Battin PH. D., Houston, Texas; Dr. I. Emich, Wien; Dr. C. Holm und R. Schmidt-Linnemann, Freiburg i. Br.; die Logopädinnen

A. Frisch und G. Hanreich, Wien; S. M. Ives, Ph. D., Florida State Univ.; Oberschulrat F. Maschka, Wien; Dipl. Log. M. Mussafia, Bruxelles; Dipl.-Log. E. Öcker, R. Kempf und R. Nesser, Linz a.d. Donau; Doz. Dr. J. Pačesová Brno; Dr. L. Striglioni, Marseille; Dr. J. Tronczynska, Lodz; Prof. C. M. Ward und C. M. Godfrey, Med. D., University of Toronto; I. Westerlund (Stockholm); Sh. L. Wirz, F.C.S.T. Edinburgh, Gr. Britain und D. Zaliouk, Tel Aviv.

DISCUSSION

Lebrun:

Les résultats de l'enquête menée par l'Association Internationale de Biophonétique confirme, dans l'ensemble, ce que l'on savait déjà de l'ordre dans lequel la plupart des enfants apprennent à réaliser les phonèmes. Ils confirment aussi l'opinion que, dans les troubles du langage, certains phonèmes sont plus que d'autres exposés à des distorsions. Toutefois, cette confirmation ne permet nullement de reconcilier la théorie phonologique de Jakobson avec la théorie physiologique de Trojan. En effet, la première nie explicitement la difficulté articuloire et assure que du point de vue physiologique tous les phonèmes sont égaux. Alors, pour laquelle des deux théories opter? Certainement pas pour celle de Jakobson, car cette théorie, comme je l'ai montré ailleurs, est infirmée par les faits. Est-ce à dire que la théorie de Trojan doit lui être substituée? Oui, à condition que l'on réussisse à montrer *expérimentalement* que la dichotomie sons pyramidaux—sons cortico—bulbaires proposée par Trojan existe dans les faits.

Dravina:

Der große Unterschied bei der prozentuellen Verteilung der Dislalie bei dem r-Laut läßt sich z. T. dadurch erklären, daß der in den Übersichtstabellen mit dem Symbol *r* bezeichnete Laut eigentlich in allen diesen Sprachen nicht gleich gesprochen wird (z.B. bereitet das Zungenspitzen-*r* für das kleine Kind größere Schwierigkeiten als das uvulare *R*).

Trojan:

Ad Lebrun: Ich begrüße die Anregung zu experimentellen Untersuchungen über die Beteiligung der cortico-bulbären und der extrapyramidalen Motorik auf das wärmste. Allerdings halte ich es für unbedingt nötig, die Enquête fortzusetzen, da uns die Erforschung der Verhältnisse bei Kindern primitiver Sprachen noch viele weitere Erkenntnisse verspricht.

DIE STATISTISCHEN UNTERLAGEN DER SPRACHLAUTCHARAKTEROLOGIE

FELIX TROJAN

Die Sprachlautcharakterologie gehört — methodisch gesehen — zur Biophonetik, ihrem Gegenstand nach ist sie zugleich ein Teil der allgemeinen Sprachcharakterologie, deren Grundzüge F. Kainz im ersten Teil des 5. Bandes seiner „Psychologie der Sprache“ entworfen hat. In der Gestalt, die ihr die Biophonetik gibt, bietet sie der Sprachwissenschaft einen grundsätzlich neuen Ansatz zur Lösung vor allem lautgeschichtlicher Probleme an. Sie räumt der objektivierenden Linguistik ein, daß das Sprachgebilde eine objektive Schöpfung des Menschen ist, führt aber die Veränderungen der Sprache, wie sie sich auf phonetischem Gebiet im Lautwandel darstellen, nicht auf Gründe im Systemgefüge des Sprachgebildes, sondern auf hormonal-vegetativ und auditiv-phänomenal begründete Gestaltungstendenzen zurück, wie sie sich beim ständigen Gebrauch des gesprochenen Wortes ergeben. Diese Tendenzen folgen prinzipiell denselben Gesetzen wie der stimmliche Ausdruck überhaupt. Die vom Sprachkollektiv geschaffene Lautfassade einer Sprache ist indes nicht als eine wesensgetreue Abbildung der Sprachgemeinschaft anzusprechen, sondern als eine intentionale, wenn auch unterbewußt entstandene Selbstrepräsentation. Sie hat grundsätzlich illusionären Charakter, mag sie für das Kollektiv auch durchaus charakteristisch sein. Sie ist in diesem Sinne einem Kunstwerk zu vergleichen. Hinter der Lautgestalt und dem Lautwandel steht also der Mensch.

Was den hormonal-vegetativ bedingten Strukturzusammenhang betrifft, so zeigt der stimmliche Ausdruck im Wechsel der Affekte, daß bei dominant parasymphathischer, trophotroper Stimmgebung das stimmhafte Element der Sprache, vorab der Vokalismus, hervortritt, bei sympathikoton-ergotroper Stimmgebung dagegen der Konsonantismus. Zugleich werden im Ausdruck des Ruheverhaltens besonders die auslautenden stimmlosen Konsonanten geschwächt, wie ja auch Offensilbigkeit im frühkindlichen Lallen und in vielen primitiven Sprachen vorherrscht. Sowohl das Verhältnis von Vokalen zu Konsonanten wie das der Offen- zur Geschlossensilbigkeit lassen sich leicht statistisch erfassen. So liegt der Prozentsatz der Vokale wie der offenen Silben im Italienischen (das wir hier und im folgenden als Beispiel heranziehen wollen) wesentlich höher als im Deutschen. Lautgeschichtlich entsprechen der Offensilbigkeit der allmähliche, fast völlige Ausfall von Konsonanten im Auslaut samt der gelegentlichen Epenthese von Vokalen im Auslaut, dem hohen Anteil der Vokale dagegen ein Trend zur Vermehrung des stimmhaften Elementes überhaupt und zur

Ausstoßung von Konsonanten, wobei sich diese letztere Tendenz freilich größtenteils dadurch verbraucht, daß sie die durch romanische Synkope entstandenen Konsonantenhäufungen reduziert. Hier läßt uns nun — und das ist das eigentliche Anliegen unserer Ausführungen — die historische Lautlehre statistisch völlig im Stich. Der Fortschritt, den G. Rohlfs „Historische Grammatik der italienischen Sprache“ gegenüber den grundlegenden Werken von W. Meyer-Lübke bedeutet, liegt im dialektologischen Bereich. Die Einführung der Statistik in die historische Lautlehre bleibt der Zukunft vorbehalten.

Im Lautbild des Italienischen wirken sich nun außer den trophotrop-regressiven Tendenzen auch gegenstrebige aus: so ein durch die Neuordnung der Quantitätsverhältnisse entstandenes Übergewicht der kurzen, offenen Silben, das auf eine niedrige Schwelle der Sensibilität hindeutet. Tatsächlich wirkt das Italienische nicht phlegmatisch, sondern lebhaft und erregt. Dazu kommt ferner eine Neigung zur Geminatio, die das Toskanische in besonderem Maße pflegt und die aus zahlreichen Quellen gespeist wird. Das rafforzamento verleiht der Sprache einen Charakter der Eindringlichkeit und der Gefühlsdichte. Sowohl das Übergewicht der kurzen Silben wie die Geminatio lassen sich statistisch erfassen, an Tonbandaufnahmen auch selektiv beobachten.

Physiologisch deutet die Verbindung der trophotropen Züge mit dem Ausdruck von Sensibilität und Gefühlsdichte auf die Wirksamkeit sexualhormonaler Faktoren. Phänomenal entsteht dadurch der Eindruck einer betont illusionären glückbeladenen Formenschönheit.

Was nun den auditiv-phänomenalen Strukturzusammenhang betrifft, so wird dieser im allgemeinen durch den Gegensatz einer Hell- und einer Dünkelsphäre beherrscht, der sich ausdrucksmäßig im Registermechanismus repräsentiert, im artikulatorischen Bereich dagegen durch die Vorderzungenvokale und die palatale Konsonanz auf der einen, die dunklen Vokale und die velaren und die labialen Konsonanten auf der anderen Seite. Die Hellsphäre hat ausdrucksmäßig Beziehungen zu lustvoller Klarheit, zum Verstandesmäßigen und zum Metaphysisch-Religiösen, aber auch zur Angst, die Dünkelsphäre zum Triebhaften und Urtümlichen. Letzten Endes spielt hier der Gegensatz von Hirnstamm- zu Hirnrindenfunktionen herein, — der Kortikalisierungsprozess, dem die Menschheit unterworfen ist.

Das Italienische schränkt die Dünkelsphäre eher ein und festigt die Hellsphäre durch die Neubildung yod-bedingter Konsonanten, wie [λ], [ɲ], [tʃ] und [dʒ].

Die Statistik des Italienischen ist durch eine dem XII. Internationalen Logopädenkongreß 1962 in Padua vorgelegte Gemeinschaftsarbeit gefördert worden. Aber auch sie würde noch einer Weiterführung im Sinne einer statistischen historischen Lautlehre des Italienischen bedürfen.

DISCUSSION

Łączkowska:

Zum sehr interessanten Vortrag von Herrn Prof. Trojan und der Aussage von Herrn Prof. Altevogt möchte ich vom naturwissenschaftlichen u. klinischen Standpunkt einige physiologische Bemerkungen, die mir sehr wichtig erscheinen, hinzufügen. Bei der Entstehung u. individuellen Bearbeitung der Sprachleistung spielt die gekoppelte Funktion des vegetativen u. des normalen Systems eine sehr wichtige Rolle. Diese Koppelung sichert die einheitliche Regelung des ganzen vegetativen Systems. Die Funktion dieser Führungszentrale (Substantia grisea centralis u. Hypophyse) ist auch weitgehend von dem Wechsel zwischen Nacht u. Tag abhängig, dessen Auswirkung über die sog. hypothalamische Wurzel des Fasciulus opticus erfolgt. Licht und Dunkelheit beeinflussen die vegetativen Abläufe in ihrem 24-Stunden-Rhythmus, auf den die parallele Umschaltung zwischen ergotroper und trophotroper Phase zurückzuführen ist. Eine Disharmonie verursacht Abänderungen des Wohlbefindens, der Reaktionsweisen und sämtlicher Leistungen, sowie auch der Sprachleistungen.

L. Novák:

Quoique les efforts du prof. Trojan et d'autres spécialistes soient très désirables, je vois, moi aussi, les difficultés méthodiques énormes contre lesquelles ils se heurtent à chaque pas. Par ex. on pourrait, de prime abord, supposer qu'il existe un *nexus causalis*, d'une part, entre le caractère vif des peuples dans la zone méditerranéenne et, d'autre part, l'accent d'intensité des langues respectives de l'union linguistique méditerranéenne dont l'existence a été proposé par moi, il y a plus de 35 ans, dans le recueil *Charisteria Guilelmo Mathesio oblata* (Prague 1932). Mais c'est justement la quantité phonologique des voyelles qui avait été auparavant partout là, avant la sémitisation progressive de ces langues à partir du IV^e siècle de notre ère, et qui n'était nullement apte pour son rythme lourd au caractère vif des peuples mentionnés. D'autre part, il est très probable qu'il existe un parallélisme surprenant entre le système clair des langues ouralo-altaïque et la vie sexuelle orientale sans scrupules quoiqu'il soit plus difficile d'en dévoiler les motivations causales précises.

Trojan:

ad Novák: Was hier bemerkt wurde, trifft das von mir Gesagte in keiner Weise. Die gemeinten Ausdruckswirkungen entziehen sich so gut wie völlig der Kontrolle des Bewußtseins. Die ethnischen Voraussetzungen spielen beim Lautwandel sicherlich eine wichtige Rolle; die lautcharakterologische Beurteilung stützt sich aber allein auf die Wirksamkeit der physiologischen Generatoren des stimmlichen Ausdrucks.

MANOMETRIC EVALUATION OF NASALITY IN CLEFT PALATE SPEECH

JADWIGA TRONCZYŃSKA*

The effects of organic deficiencies upon articulation vary considerably. A high palatal vault may influence voice, but seems to have little effect on articulation. Clefts create greater obstacles than other organic disorders because they have a profound influence on sounds dependent on substaigned pressure such as stop-plosives and fricative sounds. Some clinical and experimental observations state that complete isolation of the oral from the nasal cavity is not required for adequate articulation (Calnan, 1956, Bloomer, 1953, Kaltenborn, 1948, Milisen, 1966) but the more nearly it is achieved the less compensation will be necessary.

It is well known that the undesirable articulatory pattern in cleft palate speech develops on the basis of organic disorders. The cleft changes the pneumodynamics of oral and nasal cavity in such a direction that naso-pneumodynamics is superimposed on the oro-pneumodynamics (Mysak, 1966). This is conditioned by the too large an opening into the nasopharynx compared with the size of the opening into the oral cavity. The proper physiological conditions necessary to the good resonance relationship are inverse: The opening into the mouth between the tongue and velum must be wider than the velo-pharyngeal opening (Kaltenborn). Then the air pressure at the time of phonation will be greater in the oral than in nasal cavity.

For purposes of study the measurements of phonatory nasal pressure were performed with the aid of apparatus constructed upon principles of a manometer. The apparatus consisted of an olive joined to an elastic rubber drain connected to a metallic tube with a stop-valve. To this tube was joined a capillary tube of glass closed as a siphon. The capillary tube was fastened to the millimeter scale and filled before the measurements with a drop of alcohol solution of fuchsine. At the time of examination the olive was inserted to the nasal duct and the patient was bidden to utter isolated and syllabic phoneme sequences. The nasal air pressure during the vocalization caused the shift of stained fluid in the capillary tube. The oscillation was measured as the millimeter scale.

* From Department of Otolaryngology, Medical Academy, Łódź, Poland, Head: Prof. Aleksander Radzimiński, M. D., and from Department of Maxillo-Facial Surgery, Medical Academy, Łódź, Poland, Head: Prof. Janusz Bardach, M.D.

On the basis of 85 manometric measurements of nasality during the articulation of vowels and consonants it was demonstrated that in cleft palate speech the nasality varies in volume of expired phonatory air through the nasal ducts. The results are shown in the diagram representing the extension of nasality for all fundamental speech sounds in three groups of patients: 1. before the operation, 2. after the operation, 3. after the reeducation of speech.

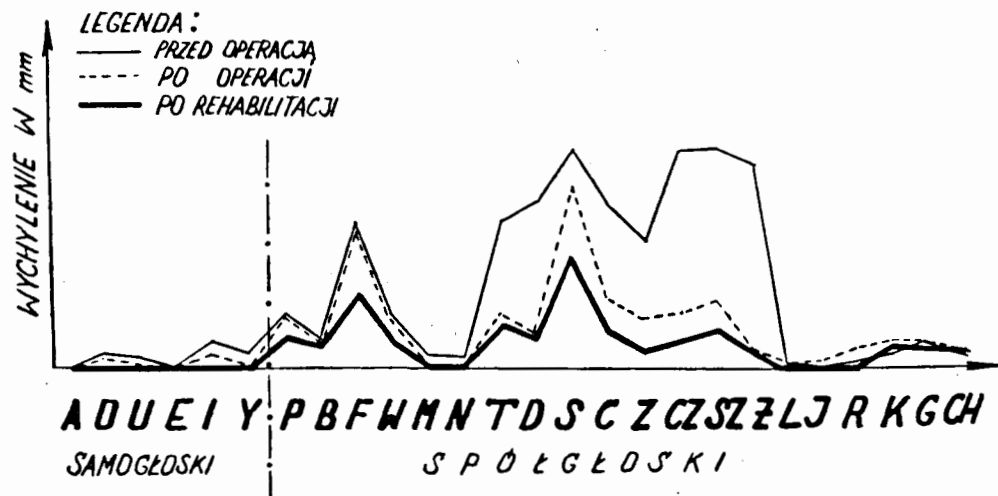


Fig. 1.

The first curve on the diagram relates to the 29 patients before the operation. The same patients were examined after the operation and the results are shown by the dotted line. Twenty seven of these patients were examined after the speech therapy and all the ratios of nasality are collected in the curve 3.

From this graphic presentation of the results is clearly seen the change of nasality in cleft speech sounds after the operation and after the speech therapy. The nasality of consonants was greatly diminished. After the reeducation if even there were slight traces of nasality in manometric measurements in general it was not heard in speech.

From our investigation some observations of phonatory value result. It was evident that in the majority of our measurements the greatest ratio of nasality was stated in the production of "F" and "S". It can be concluded that the greatest amount of oral breath pressure is needed in the production of these sounds of speech which require the most narrow passage of the air through the half-closed mouth. The velopharyngeal gap in cleft palate patients is too wide for the production of sufficiently great oral pressure. Thus it can be seen that in speech production always and in the articulation of all sounds of speech the same proportions between the openings of two resonance chambers i.e. nasal and oral cavity must be kept. Always the oral opening must be larger than the nasopharyngeal one.

Of great importance in voice rehabilitation in cleft palate patients is auditory monitoring. The comparison of two groups of patients: after the operation and after the speech therapy show that the degree of nasality did not change so much as did the intelligibility of speech, which greatly improved after the rehabilitation program.

However, our manometric measurements of nasality stress the significant role of the relationships between two resonatory chambers i.e. nasal and oral cavity and show that the efforts of surgery should aim to enlarge an oral port and to reduce the nasopharyngeal opening.

BIBLIOGRAPHY

1. Calnan, J.: Movements of the Soft Palate. *Brit. J. Plastic Surg.* 5 1953.
2. Bloomer, H. H.: Observations on Palatopharyngeal Movements in Speech and Deglutition. *J. Speech Hear. Dis.* 18, 1953, 230—246.
3. Kaltenborn, A. L., cit. by Rieber and Brubaker, *Speech Pathology*, Amsterdam, 1966, p. 329—330.
4. Milisen R.: Articulatory Problems, chapt. 7 in "Speech Pathology" by Rieber R. W. and Brubaker R. S., 1966, Amsterdam, p. 137—149.
5. Mysak, E. D.: Pneumodynamics as a Factor in Cleft Palate Speech. *Pl. reconstr. Surg.*, 28, 1961, 588—591.

DISCUSSION

Vrtička:

It is necessary to distinguish "nasality" as an acoustical parameter from "nasal escape" as an aerodynamic one. In many cases, these two values do not exactly correspond.

VOKALQUANTITÄT UND SILBE IM DEUTSCHEN

PAVEL TROST

Die Frage Silbenschnitt (Anschlußart) im Deutschen wurde oft behandelt, aber nicht entschieden. Die Ansicht, daß der stark geschnittene Akzent (feste Anschluß) auf Oszillogrammen sichtbar wird, und zwar als steiler Abfall der Vokalkurve, begegnete auf dem Kongreß von Helsinki vorsichtigem Zweifel. Dabei wurde der subjektive (auditive) Eindruck des festen Anschlusses nicht in Zweifel gezogen. Doch es blieb bei der Vermutung, daß dieser Eindruck vielleicht nur auf Kürze und Ungespanntheit des Vokals beruht. Aber so vielen Sprachen mit vokalischer Quantität wird kein geschnittener Akzent nachgesagt, obwohl sich auch in diesen Sprachen mit dem Quantitätsunterschied ein mehr oder weniger deutlicher Unterschied in der Qualität (Gespanntheit) verbindet.

Nicht berücksichtigt wurde in der Diskussion von Helsinki der Gesichtspunkt der Silbenbildung. Als Sievers die Lehre vom Silbenakzent entwickelte, hatte er auch ausgeführt, daß sich Silbenschnitt und Silbenbildung miteinander berühren, doch ohne daß seiner Meinung nach der Silbenschnitt von der Silbentrennung ohne weiteres abhängig ist. Nach Sievers liegt in der deutschen Bühnensprache in zweisilbigen Wörtern mit kurzem betonten Vokal der ersten Silbe wie *fassen*, *Kummer*, *alle*, die Silbengrenze (im Einklang mit der Rechtschreibung) zweifelsohne im Konsonanten: dieser scheint zu beiden Silben zu gehören, weil in seiner Artikulation keine Diskontinuität der Druckstärke eintritt. Andere Phonetiker sind jetzt sogar der Meinung, daß hier die Silbengrenze nach dem Konsonanten liegt. Wenn in diesen Fällen die Silbengrenze in oder nach dem Konsonanten liegt, so hat das zu bedeuten, daß betonte kurze Vokale im Deutschen nur in geschlossenen Silben vorkommen. Darin liegt dann ein besonderer Zug, der das Deutsche von anderen Sprachen mit vokalischer Quantität unterscheidet. Wenn das stimmt, so läßt sich der Silbenschnitt grundsätzlich auf die Silbenbildung zurückführen. Aber der Ansatz der Silbengrenze im Konsonanten ist bisher durch Oszillogramme oder Sonagramme nicht bestätigt.

Der Ansatz der Silbengrenze im Konsonanten ist jedoch von der Sprachstruktur aus begründet. Wenn einsilbige Wörter das Strukturmodell der betonten Silben liefern, so kennt die deutsche Hochsprache keine betonten Silben, die mit kurzem Vokal schließen: im Deutschen gehen Einsilbler entweder auf langen Vokal oder konsonantisch aus. Schon bei Trubetzkoy wurde der Ansatz der Silbenschnittkorrelation im Deutschen und in den anderen neugermanischen Sprachen vom phonologischen

Standpunkt aus begründet. Trubetzkoy hatte auch vokalische Quantität und Silbenschnitt zueinander in Gegensatz gestellt: bei der Quantitätskorrelation ist die Länge merkmalshaft, bei der Silbenschnittkorrelation ist dagegen die Kürze mit dem merkmalshaften Glied verbunden und Länge des Vokals erscheint in der Aufhebungsstellung.

Es fragt sich also: Muß der auditive Eindruck des festen Anschlusses vom akustischen Signal oder allein von der Sprachstruktur aus erklärt werden?

Der auditive Eindruck ist selbst nicht unproblematisch. Längst wird darauf hingewiesen, daß der stark geschnittene Akzent in bestimmten Bereichen der deutschen Sprache überhaupt fehlt. Unsicherheit besteht darüber, ob in der Hochsprache auch in Fällen wie *Halm, halten* fester Anschluß zu hören ist. Nun unterscheiden sich solche Fälle wie *Halm, halten* von anderen wie *betten, bitten* in der Silbenstruktur: in *Halm, halten* ist der konsonantische Silbenschluß auch ohne festen Anschluß gesichert. Wenn daher der ausgeprägte stark geschnittene Akzent in der Regel auf Fälle wie *betten, bitten* beschränkt ist, so erweist er sich als Äquivalent von konsonantischem Silbenschluß. Der feste Anschluß kann aufgefaßt werden als ein zu vokalischer Kürze hinzutretendes Element der Silbenstruktur, gebunden an einen vorhergehenden kurzen betonten Vokal mit einem folgenden einfachen intervokalischen Konsonanten.

Es ist ein banaler Vorgang, daß ein Merkmal, das ein anderes begleitet, schließlich dessen Funktion übernimmt. Die vokalische Quantität hat im Deutschen im Laufe der Entwicklung eine Minderung erfahren. Am Anfang steht die Beschränkung des vokalischen Quantitätsgegensatzes auf betonte Silben. Am Ende steht ein Zustand wie in den bairischen Mundarten, wo die Vokalquantität durchaus an den Stärkegegensatz des nachfolgenden Konsonanten gebunden ist. Hier kann man die Vokalquantität als nichtrelevant und abhängig vom relevanten Stärkegegensatz der Konsonanten ansehen. D. h. die distinktive Funktion der Vokalquantität ist ohne Einschränkung auf Konsonanten übergegangen. Ebenso begleitet nicht nur der Silbenschnitt die Vokalquantität, sondern kann sie auch supplieren. Genetisch ist übrigens der Silbenschnitt mit der konsonantischen Geminierungskorrelation verknüpft, die mit dem sog. neuhochdeutschen Quantitätsausgleich die phonologische Relevanz einbüßte. Der Quantitätsausgleich drang bekanntlich nicht in den alemannischen Mundarten durch; dazu stimmt das Fehlen des stark geschnittenen Akzents im „Südwestdeutschen“.

Vom phonetischen Standpunkt aus hat man die Lehre vom Silbenschnitt als Non-sense bezeichnet, weil ein Vokal nicht unmittelbar von einem Konsonanten abgelöst werden kann. Aber der Übergang vom Vokal zum Konsonanten läßt bestimmte Modifikationen zu. Die phonetische Analyse hat unabhängig von der phonologischen Betrachtung festzustellen, was im akustischen Signal enthalten und was nicht enthalten ist. Doch wenn die phonologische Betrachtung des festen Anschlusses auf den Konsonanten hinweist, so ist das immerhin ein Hinweis für die phonetische Untersuchung, im Konsonanten nach dem Merkmal des festen Anschlusses zu suchen. Wenn sich dann wirklich im akustischen Signal kein Korrelat findet, so bleibt nichts

übrig, als den auditiven Eindruck des stark geschnittenen Akzents (bestätigt durch den Vergleich der vokalischen Kürze im Deutschen und etwa im Tschechischen) allein von der Sprachstruktur aus zu erklären. Ich stelle mir die Erklärung in diesem Sinne vor: Wenn sonst nicht der Anglitt, sondern nur der Abglitt des Konsonanten sprachlich relevant ist, d. h. gesondert wahrgenommen wird, so in dieser Stellung auch der Anglitt, und zwar kraft des Gesetzes vom Silbenschluß im Deutschen.

DISCUSSION

Fischer-Jørgensen:

Die Frage des Silbenschnitts ist sehr heikel. Ich habe vor Jahren den Intensitätsverlauf von langen und kurzen Vokalen im Deutschen untersucht, ohne einen Unterschied finden zu können. Ich möchte aber besonders auf die Untersuchungen von A. L. Fliflet hinweisen. Er hat auf Grundlage von Abhörversuchen, wo auf dem Tonband z. B. der Anfang des Vokals oder das Ende des folgenden Konsonanten abgeschnitten waren, nachgewiesen, daß die Beurteilung des Silbenschnitts (durch Gruppen schwedischer und dänischer Phonetiker) von der relativen Dauer von Vokal und Konsonant weitgehend beeinflußt war.

Heike:

Der Vortragende äußerte die Ansicht, daß der subjektive Eindruck des „festen Anschlusses“ im Falle mangelnder physikalischer Korrelate nur durch die Sprachstruktur des Deutschen bestimmt sein kann. Subjektive Eindrücke können Linguisten und naive Hörer haben. Nur naive Hörer jedoch können von der Sprachstruktur in ihren Urteilen bestimmt sein. Die Terme „fester Anschluß“ etc. stammen jedoch von Linguisten. Naive Hörer hören bspw. *biete* oder *bitte*, aber nicht losen oder festen Anschluß.

Lehiste:

I have not been able to identify any acoustic evidence for “festen und losen Anschluß”. However, I have some recent evidence for syllable boundaries within a consonant. In an electromyographic study of the orbicularis oris muscle, I found two peaks of activity for cases in which an intervocalic /P/ was a geminate, and only one peak when the syllable boundary was judged (on linguistic and subjective auditory grounds) to lie before the /P/. The test material consisted of estonian words *TABA, TAPA, TAPPA*.

Martens:

Nur ein Vorschlag zur Vermeidung von Mißverständnissen: Man sollte vielleicht besser nicht nur von „kurzen“ Vokalen sprechen, sondern genauer bezeichnen, was gemeint ist; ob ein „kurzer, offener, ungespannter Vokal“ oder ein „kurzer, geschlossener, gespannter Vokal“. Es ist nicht richtig zu sagen: „kurze Vokale gibt es nur in geschlossenen Silben“ und „es gibt keine Silben, die auf kurzen Vokal ausgehen“. Es gibt solche Silben im Deutschen eben doch. Aber mit *kurzem, geschlossenem, gespanntem* Vokal: z. B. *Juni, Juli, Auto, Kino, Akku*. Es sei denn, man rechnet diese Distribution nicht zu dem zentralen, sondern zu den peripheren Systemen; aber dann bleiben immer noch Wörter (wie z. B. vielleicht [*f i ' l a i ç t*]) mit Sprechsilbengrenze nach dem kurzen, geschlossenen, gespannten Vokal.

THE PERCEPTION AND COMMON MISPERCEPTION OF INFANT PRE-SPEECH

H. M. TRUBY*

INTRODUCTION

In an attempt at harmony with the announced central theme of this Congress, the following observations are here offered for consideration. Acoustic signal, after all, is tangible. Its perception is most demonstrably *not!* The specification, acoustic speech signal, *can* be held to the tangible, but is not usually so restricted, since most evaluation is based on phonemic and higher level linguistic principles than the uninvolved, purely physical (to include biophysical) analysis entails. *Phonological* implications, once admitted, alter the course of objectivity... quite beyond recognition, it can fairly be added. Were this not so, the dichotomy implied by the two rubrics, *phonetics* and *phonetic sciences* might never have come to be entertained. Speech is the concern of phonetics—the speech signal is one of the concerns of the phonetic sciences. It is impractical—if not impracticable—for the phonetician to avoid dialectology or phonemics by whatever names. Various departments of the phonetic sciences are, interrelatedly, directed at various physical—to comprehend biophysical—areas. The measurements which define these departmental interests should depend in no way upon perceptual proclivities, and one of the most obvious and accessible and expounded of these measures is the acoustic. It is troublesome—but of little consequence—that the disciplines of phonetics, of the phonetic sciences, and of phonology are rarely treated in mutual exclusion. It is, however, of considerable consequence that *physical* inferences are quite often drawn from *psychophysical* exercises. This paper is directed at illustrating the consequence of such misappropriation.

DISCUSSION

The developmental sector, “infant pre-speech”, is directly the concern of phonetics. More has been miswritten about developmental speech than a dozen papers of this scope could hope to offset. By selecting the *pre*-speech sector, linguistic criteria are by definition irrelevant. By admitting perception—and misperception—the full

* Communication Research Institute, Coconut Grove, Florida. (Now, Professor of Pediatrics and Linguistics, University of Miami, Coral Gables, Florida.)

complement of psychophysical vectors and parameters may be treated. Let us now consider in what ways these conflicts breed invalidity.

When acoustic signal is mentioned, sound is explicit. Where there is sound, there is motion to which that sound relates... from which it derives physically.¹ In the human, where there is conventional speech-sound motion, there is neuromuscular transmission, purely neural pregeneration, and something that may as well be referred to as conceptual *pro*generation. All of this biophysical activity is as deserving of the name *speech* as is the resultant acoustic activity, and the same holds for the *pro*generating conceptualizing process.² Yet, insistence on such definition is clearly moot. [The roles of feedback or of the listener or receiver are purposefully avoided here.] Our present objective is the derivative relationship of the acoustic signal to its generatory motion.

Sound is enigmatically elusive insofar as memory registration is concerned.³ There simply is no way to describe most sound continua. Speech sound is, to a degree, one of the rare exceptions, but even with speech sound, not only is the pattern-conditioning literally indescribable, but there is much of the speech-sound signal which is neither recoverable nor registerable. Anyone who has ever attempted to re-create or just plain remember the voice of an acquaintance or a family member or whomever finds it impossible to proceed. Which criteria can be used? True, certain timing and emphasis and voice quality features can be more or less indicated, but on the whole, a faithful physical recounting is not to be had. Yet it cannot be denied that the recognition aspect, for example, is clearly no less real than the identification with-certain-phonemes aspect. How does such a situation come to be? The answer is in the pattern-matching criteria directly, society having provided the incentive of communication to speech-element identification generally, while the phenomenon of speaker-recognition has remained on the whole a more individual or small intimate group consideration, public figures perhaps to the contrary. Yet one sound aspect is as physically real as the other—voices can certainly be identified by those familiar with the details of their acoustic manifestation, and if they are permanently recorded in some audible form, they are permanently available for re-identification. Otherwise, there is *no* physical medium form in which the details implicating the inherent acoustic features may be described differentially—there is no sound-symbol system keyed to voice quality minutiae. Alphabets are basically speech-sound-symbol systems, but no one would for a moment attempt to transcribe voice-quality differences with conventional alphabet symbols. Nor would the phonetician have any

more success with any so-called phonetic alphabet. Yet, with complete disregard for the irrelevance of it, more than a few students of speech have transcribed—in terms of speech-sound symbolism—a wide variety of events to which their only analysis access was auditory. The case in point is the catalyst for this report.

Physiological phonetics has traditionally made an issue of the fact that speech is an overlaid function of respiration, deglutition, even mastication, and the like. However, the focus has quite understandably always been upon articulation, with some spillover into the aspects of source generation, intonation, emphasis, and timing, for instance. Whenever subphonemic details are marked, they relate in turn to articulatory features primarily and associated physiological aspects secondarily, but all with the *basis* in mature physiology and mature performance. Consider, if you will, the all-important area of newborn-infant cry-sound analysis. Various combinations of direct observation, X-ray motion-picture photography, sound-spectrographic records, and the like reveal that the infant uses a wide variety of compensatory articulatory and sound-generation mechanisms to accomplish what seem to be “linguistically same” vocal productions. The representation of these performances using conventional phonetic transcription (the basis of which is, as stated, in adult physiology and anatomy) is demonstrably *inappropriate* to the task of *pre*-speech-sound evaluation. From the purely physical point-of-view, cry-sound sequences have “vowels” and “consonants”, and nasalization, and voicelessness, and they differ from each other in their spectral envelopes, and durations, and intonational effects as well, *but*, cry sound should certainly not be analyzed as though it were speech. In fact, much damage has been done by observers who have assumed that when they—the observers—perceive the presence of a particular sound image in an infant emission, they should indicate in their transcription and in their evaluation the articulatory mechanism conventionally associated with that image. Nothing could be further from the facts, which are, essentially, that differing articulations will not produce the phonetically same oscillogram or sound spectrogram, no matter how phonemically alike these sounds may appear to be! Phonemics simply has no place in the analysis of the non-speech signal, and extreme care is invited wherever compensatory mechanisms are apt to be general (as in pathology) or wherever analogs are employed (as in so-called “speech synthesis”).

The archives of this researcher contain thousands of recorded cry-sound sequences, many of which were made with correlated cineradiographic monitoring. Sound spectrograms of the tape-recorded events, as the recorded events themselves, carry acoustic cues which, for the native English speaker, would seem to correspond to conventional speech-sound sequences. For instance, hundreds of sequences would appear to be describable as [mãe:], (see A of figure) yet, inspection of the cineradiographic print of any of these sound productions would verify that the infant's lips never even approximated bilabial occlusion and, in fact, that his mouth was rather wide open, the tongue suspended under tension and concave in more or less low central position, and that all relevant articulation—that is, all articulation resulting in the

¹ Cf. Truby and Lind, “Cry Sounds of the Newborn Infant”, *Acta Pædiatrica* 163, 1965, p. 9, and see Footnote 3 herewith.

² Truby, “A Definition of Speech-sound Analysis, “Speech Synthesis”, and Speech”, André Martinet Festschrift, *LINGUISTIC STUDIES*, Clowes, N. Y. — London, 1968.

³ Truby, Bosma, Lind, *Newborn Infant Cry*, *Acta Pædiatrica Scandinavica* 163, 1965, p. 17. (first published as Communication Research Institute Scientific Report No. 0260, 1969).

perceived bilabial nasal continuant—took place in the mesopharynx and consisted of very evident approximation of mesopharyngeal wall to tongue-root! Practically all egressive, infant cry sound is characterized by the tongue position described, insofar as the most vowel-like portion is concerned, and practically all activity which could be construed as articulation takes place in the mesopharyngeal area as illustrated above. Thus (see B of figure), there is no alveolarity in articulatory evidence during a cry-sound sequence heard as [nã:] (“dentality” being ruled out in the absence of teeth!), and so on throughout any inventory correspondences imagined by a listener (see, for example, C of figure). Medical doctors unacquainted with basic phonetic principles can be excused for their misvaluations (though not forgiven for not seeking out professional counsel), but it seems a shame that reputable phoneticians-by-portfolio must also be included among the culpable. The, I hope unforgettable, lesson-to-be-learned is: “Do not assume an articulation or other phonetic feature on the basis of acoustic evidence alone!” The acoustic facts may well be objective and valid and reliable, but phonetic identifications by observers are subconsciously and almost ineradicably based on habitual adult articulatory performance and have little or no reference to sound-producing mechanisms per se.

All infant vocalization leads, in some sense, to the ultimate stabilization, essentially, of speech-performance patterns, and attempts to account for the unfolding details of these developmental pathways have proved intriguing to many professionals and non-professionals alike and will continue to prove so, but there is good reason to proceed cautiously as suggested above. At certain later stages of what may be termed “pre-speech” performance, there is relevance to phonetic evaluation in terms of adult linguistic-signal processing, but at earlier stages, such evaluations are generally, if not usually, highly irrelevant. It is not likely that technological instrumentation will be developed to account for the infinite variety of phonetic interpretations, and all premature evaluations (by investigators, and in the senses indicated) are by definition suspect.

That which has been offered as regards newborn-infant sound production could, with impunity, be extended, as every linguist among us will corroborate, to any and all *inter-language* considerations, and the International Phonetic Association made itself irrefutably clear on this point right from its beginnings. The significance of the IPA assertion that visual symbols are only by convention referable to particular vocal events has been essentially lost in the last important area of misappropriation which this report shall treat, and that is the area which has come to be known as “speech synthesis”. Since what has to be said in this regard is at once obvious in the light of the material discussed above, it must suffice to say that the *product*—the signal output—of the synthesizer might more aptly be termed *pseudospeech*, and the implications will be at once evident to every member of this Congress. Or certainly *should* be! Machines simply do not have teeth nor lips nor tongues nor the like, and the sort of transcription which applies to their output is a bird of a different breed. Admittedly the output is audible, and insofar as the genius of the programmer is

concerned, the machine program is related to human programs. At that point the analogy ends.

In order to approach the conclusion of this paper with a highly relevant and serious consideration, may all of its consumers be reminded that the success or failure of treatment in pathology or the validity of research rest with the phonetic diagnosis. An infant with loss of neuromuscular coherence in the tongue-blade area whose output is evaluated by the phonetician in terms of apico-alveolar performance, has little chance of being exposed to relevant therapy until the phonetician’s misdiagnosis is discovered. It is not enough to have “a highly trained ear”, nor, in the case of sound spectrograms, a “skilled eye”. Nothing can take the place of an immaculately plenary consideration of all the possibilities. It is to be hoped that the suggestions offered here will assist the phonetician in capitalizing on the full store of information his special training has equipped him to uncover.

CONCLUSIONS

Conventional phonetic transcription,⁴ based as it is on mature-speaker anatomical, articulatory, and associated speech-feature particulars, is inappropriate to the task of describing *immature* (in whatever sense) sound performance,⁵ which though emanating, as speech, from the upper respiratory tract, may be seen to be more remotely constituted *from* speech than is immediately apparent. This does not mean to infer that there is not much of great interest to be considered as regards the developing program of employment of varying (with maturation) physiological mechanisms from speech-sound to speech-sound and from individual to individual. However, as suggested in the opening thoughts of this paper, the acoustic signal is, in a sense, tangible—the perception of that signal, being a psychophysical operation, is at once arbitrary, illusionary, and based, at best, on an unstable inventory of phonological bias. It is thus best to proceed with caution where the objective is physical description. *Communication* is one thing—the measurement of *modi* and *media* is quite another.

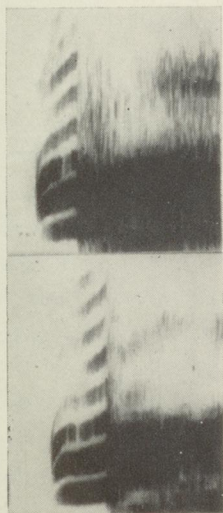
⁴ Cf. Truby, “Pleniphonetic Transcription in Phonetic Analysis”, *Preprints* (Harvard-MIT, 1962) and *Proceedings* (Mouton, 1964) of IX International Congress of Linguists and the “phonette” in *Acoustico-Cineradiographic Analysis Considerations, Acta Radiologica* 182, Stockholm, 1959.

⁵ Truby, “Cry Sounds of the Newborn Infant” (op. cit.), pp. 17—18, 57. “Language and Dolphins”, Xth International Congress of Linguists, Bucharest, 1967.

DISCUSSION

Slama-Cazacu:

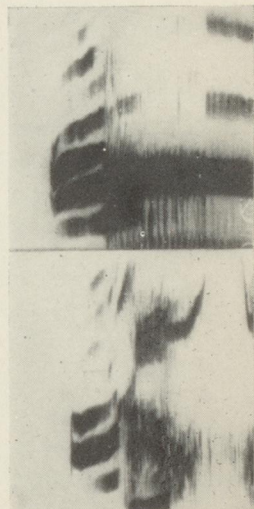
Les deux communications (Mme Sedláčková et M. Truby) touchent un problème qui commence à être exploré avec des procédés objectifs. Des recherches de ce genre sont utiles non seulement pour le problème du développement du langage, mais aussi pour la phonétique générale. Elles m'ont intéressée aussi d'un point de vue personnel: nous avons initié une recherche, en collaboration internationale, concernant la formation du système phonématique chez l'enfant et des recherches de détail, si minutieuses, comme celle de Mme Sedláčková, nous seront très utiles. Je voudrais souligner aussi la nécessité de réaliser — par une coopération — un système international de transcription phonétique adéquat au pré-langage enfantin.



A



B



C

- A — Sound spectrograms of newborn-infant cry-sound initiations resembling [mã:], but see related text!
- B — Sound spectrograms of newborn-infant cry-sound initiations resembling [nã:], but see related text!
- C — Sound spectrograms of newborn-infant cry-sound initiations resembling [ŋã:], similarly to above cited circumstances.

ÜBER DIE GESTALTKRITERIEN AKUSTISCHER SPRACHSIGNALE

WALTER TSCHESCHNER*

Bei der Untersuchung der akustischen Zeichenerkennung durch organische Systeme, insbesondere der Sprachsignale, erhebt sich die Frage, welche Signaleigenschaften werden dem psycho-physischen Erkennungsvorgang zugrunde gelegt.

Beschränkt man sich auf die akustischen Signaleigenschaften, die die für die Artikulationseinstellung des organischen Senders typische Sendefunktion beschreiben, so reichen die Eigenwerte des Resonanztraktes bzw. die Formantdaten aus, um die semantische Funktion eindeutig zu erfassen. Die Formanten¹ können aus der Konfiguration des Ansatzrohres, also des Sendesystems, genau berechnet werden (1). Ihre exaktere Erfassung aus dem Sendespektrum gelingt wegen dem diskontinuierlichen Anregungsspektrum bei stimmhaften Lauten nur über Näherungsverfahren (2). Die besondere Bedeutung der Formanten ist dadurch gekennzeichnet, daß bei ihrer hinreichend genauen Bestimmung nach Frequenz, Bandbreite und Amplitude drei Formanten genügen, die Zuordnung Zeichenträger — elementares Sprachzeichen (Artikulationsstellung) zu finden.

Für den Entwurf technischer Spracherkennungs- und Verarbeitungseinrichtungen ist es nun von großer Bedeutung zu wissen, welche Kenngrößen der adäquate organische Empfänger aus dem Signalkurzzeitspektrum extrahiert und der Erkennung zugrunde legt. Wie bereits früher mit Hilfe der subjektiven Analyse — durchgeführte Untersuchungen zeigten, stützt sich der organische Empfänger nur mittelbar auf die Extrahierung von Formantdaten (3). Zur Erfassung der primären empfangsseitigen Kenndaten von Lautspektren wurden in weiteren Versuchen u. a. auch synthetische Lautspektren verwendet.

Über die dabei gewonnenen Ergebnisse soll hier berichtet werden.

Zu den Versuchen wurde ein handgesteuerter Synthesator (Bild 1) benutzt, der nach dem Prinzip der Sprachdecodierung eines Kanalvocoders (4) arbeitet. Das System besitzt 12 parallel geschaltete Netzwerke, deren Durchlaßbereiche und Frequenzstaffelung etwa der Frequenzgruppenbildung entsprechen:

* VEB Elektronische Rechenmaschinen.

¹ Nach Fant (1) ist es zweckmäßig, den Terminus Formant nur auf die Dämpfungsnullstellen des Ansatzrohres anzuwenden.

0 — 0,25 — 0,35 — 0,45 — 0,6 — 0,85 — 1,15 — 1,45 — 1,75 — 2,1 — 2,5 —
— 2,95 — 3,5 kHz

Das Spektrum jedes Reizes wurde über Steckbuchsen der steuernden Widerstandsmatrix (Bild 1) vorgegeben. Die Auslösung wurde von Hand vorgenommen, wobei ein Zeitfunktionsgenerator mit einer kontinuierlich verlaufenden Zeithülle von 200 ms Länge freigegeben wurde.

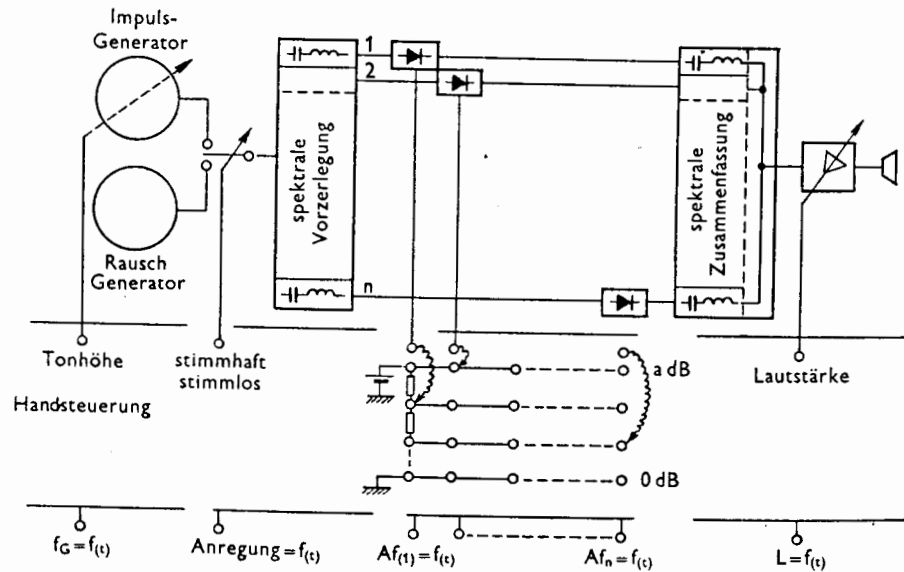


Abb. 1. Schema der Sprachsignalsynthese.

Als Anregungsparameter wurden verwendet:

Stimmhaft: Dreieckfunktion — 6 dB/Oktave ausgeglichen auf 0 dB/Okt, $f_G = 110$ Hz, Grundfrequenzmodulation zur Erhöhung der Natürlichkeit $f_{G \text{ Hub}} = +20$ Hz, $f_{\text{mod}} = 0,2$ Hz

Stimmlos: Weißes Rauschen 0 dB/Oktave.

Je 50 Reize, von denen jeder 4 bis 5mal vorkam, wurden in einer bestimmten zeitlichen Reihenfolge auf Band zusammengestellt und über dynamische Meßhörer bei einer mittleren Lautstärke von 80 phon von 5 Hörern beurteilt. Nach Vorgabe eines nichtbewerteten Versuches konnte jeder Reiz, sofern akzeptabel, mit einem geläufigen Lautnamen belegt werden. Es ergab sich jedoch, daß die Zahl der nicht treffbaren Zuordnungen außerordentlich gering waren.

Im 1. Versuch wurde nur jeweils ein Kanal bei stimmhafter Anregung geöffnet. Dabei wurde darauf geachtet, daß die zeitliche Aufeinanderfolge ähnlicher Reize vermieden wurde. Beim 2. Versuch wurden die gleichen Einstellungen mit stimmloser Anregung wiederholt. Im 3. Versuch wurden bei jedem Reiz der in Bild 2 markierte

Kanal mit voller Amplitude geöffnet und entsprechend dem 1. Versuch ein weiterer Kanal bei stimmhafter Anregung zugeschaltet.

Das Ergebnis der Untersuchung in Bild 2 enthält den Prozentsatz der pro Reiz getroffenen Zuordnungen (5).

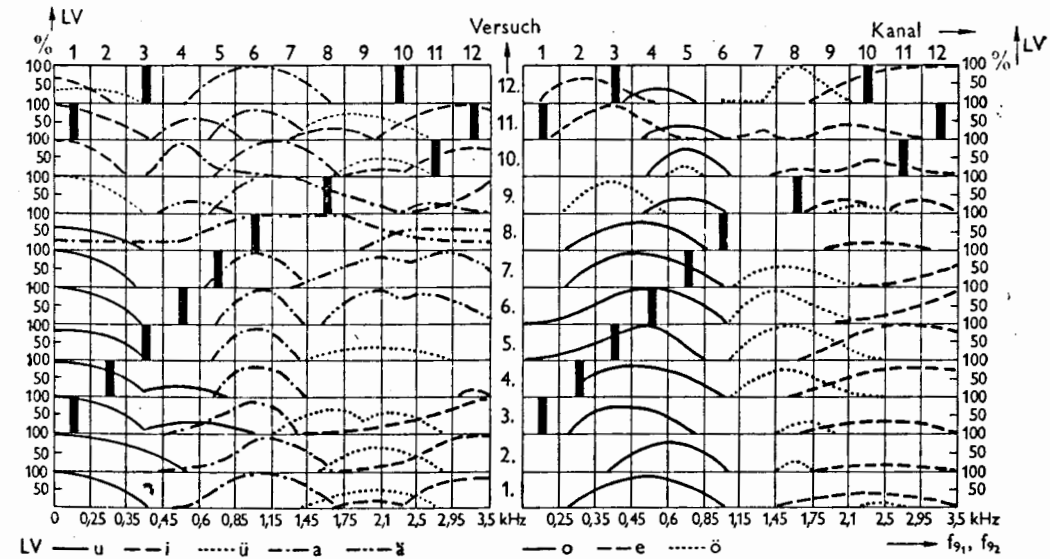


Abb. 2. Die Lautverständlichkeit synthetischer Spektralkonfigurationen.

Wird die Formanthypothese zur Deutung der Versuchsergebnisse herangezogen, dann müßten sich die Hörerurteile nach der folgenden Formantentabelle für deutsche Vokale richten.

	$F_{1\text{kHz}}$	$F_{2\text{kHz}}$	$F_{3\text{kHz}}$
u	0,25	0,6	
o	0,4	0,8	
ä	0,6	1,2	2,2
a	0,85	1,2	2,5
ö	0,4	1,8	2,4
e	0,4	2,4	3,25
ü	0,35	2	2,8
i	0,25	2,6	3,5

Demzufolge dürften im 3. Versuch nur die Lautvalenzen [u], [ü] und [i] im 5. Versuch nur [o], [ö] und [e] und im 6. Versuch nur [u] und [ä] in Erscheinung treten. Nach Bild 2 zeigen sich jedoch andere Ergebnisse.

Im 3. Versuch werden z. B. teils unsichere [ä] 5 %, [ö] 25 %, [e] 30 %, teils sichere [ü] 65 %, [o] 80 %, teils sehr sichere [a] 90 %, [i] 95 % [u] 100 % Zuordnungen gefunden. Diese Werte stellen Maximalwerte dar, die rechts und links davon auf der Kanalskala nach Null verlaufen. Entsprechende Verhältnisse sind in allen Versuchen zu beobachten.

In der gewählten Versuchsanordnung, bei der den Hörern Relativentscheidungen möglich sind, werden sehr sichere Vokal-Zuordnungen (> 90 %) getroffen, wenn folgende Kanäle geöffnet werden.

Kanal

u	1	2+1	4+1	5+1		1							
o	4	3+4	4+3	5+3		4							
ä	4+9	5+11				5+10							
a	6	1+6	2+6	3+6	4+6	5+6	6+7	6+8	8+6	8+7	11+7	3+10+6	6
ö	3+8	3+10+8											3+8
e	3+11	3+10+11	1+12+3	3+10+12									3+11
ü	8+1												1+8
i	12	1+12	11+1	1+12+11									1+11/12

Es ist zu erkennen, daß zwar für die Kanalmuster, die näherungsweise der oben angegebenen Formantkonfiguration entsprechen, weitgehend die zugehörigen Lautnamen gefunden werden, für die verschiedensten anderen Kanalmuster werden aber ebenfalls sehr sichere Zuordnungen angegeben. Somit kann das Formantprinzip nicht der Erkennung zugrundeliegen. Wird dagegen hier zur Beschreibung des Erkennungsprozesses die Gestalt und nicht einzelne diskrete Punkte der Spektralhülle herangezogen, so kann der Versuch befriedigend gedeutet werden.

Sieht man von den zusätzlichen Einflußfaktoren, wie Adaption, Folgekontrast usw. ab, so können aus diesem Versuch, ergänzt durch Einzelversuche, etwa die in Bild 3 enthaltenen Gestaltkonfigurationen der Spektralhülle der Erkennung deutscher Vokale zugrundegelegt werden. Diese Figuren, die man sich als Transformation der Auslenkung der Basilmembran vorzustellen hat, werden durch kettenlinienförmige Merkmalschwellen begrenzt. Als typisch sind die Formmerkmale und erst in zweiter Linie die Frequenzlage anzusehen. Formmerkmale sind z. B.:

- [u] Frequenzgruppenbreite Erregung niederfrequenter Lage. Besonderes Merkmal niederfrequente Anteile unter dem Maxima, mindestens um 40 dB gedämpft.
- [a] Zentrale maximale Erregung, ca. 3 Frequenzgruppen breit.
- [i] Ausgeprägtes Minimum, dessen Sohle mindestens 25 dB unter den Maxima liegen muß.

Ist ein solches Formmerkmal Bestandteil eines Kanalmusters, dann wird der entsprechende Laut zugeordnet, wobei offenbar in der Erkennungslogik noch eine hierarchische Rangordnung wirksam wird (6).

Somit kann einem akustischen Reiz analog einem optischen Reiz eine Gestaltkonfiguration und eine Färbung (Grundfrequenz) zugeordnet werden.

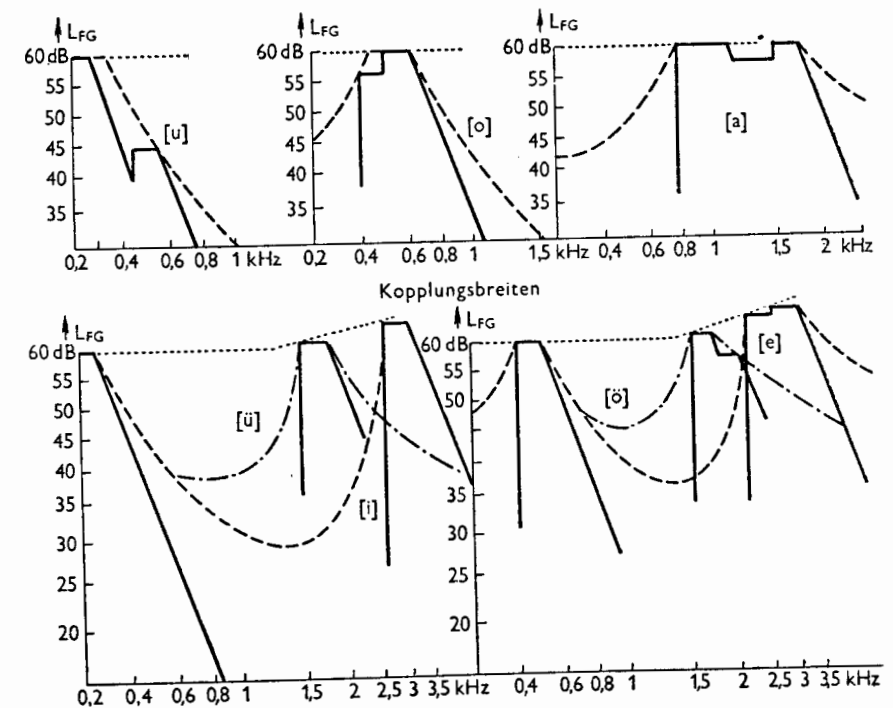


Abb. 3. Die zu einer Lautvalenzklasse eines deutschen Vokales gehörige empfundene Gestaltfunktion ermittelt durch Syntheseuntersuchungen.

- dominierendes Gestaltmerkmal.
- Merkmalschwelle.
- Größe der spezifischen Lautheit bei konstantem Frequenzpegel.

LITERATUR

1. Fant, G.: *Acoustic theory of speech production*. Verlag Mouton Co. 1960, 's-Gravenhage.
2. Lindblom, B.: Accuracy and limitation of sonagraph—measurements. *Proc. IV. Phonetik Kongress Helsinki 1961*, S. 188—200.
3. Tscheschner, W.: Analyse der deutschen Sprache unter besonderer Berücksichtigung der nichtstationären Vorgänge (Teil 4). *Zeitschr. f. Phonetik und Kommunikationswissensch.* 19 (1966), S. 141—201.
4. Krocke, E.: Aufbau und Untersuchung eines Übertragungssystems für synthetische Sprache. *Wiss. Zeitschr. der TH. Dresden* 6 (1956/1957) H. 4, S. 757—776.
5. Tscheschner, W.: Die Sprache und ihre Erkennung. Habilschrift (unveröffentl.). TU Dresden 1967.
6. Tscheschner, W.: Ein Beitrag zur subjektiven Erkennung akustischer Sprachlaute. *IV. Ak. Konferenz Budapest 1967* (erscheint im Tagungsbericht).

DISCUSSION

Fant:

Your results appear to support a two-formant model of speech perception in the sense of two major spectral stimulus regions.

ZUR AUDITIVEN INTERPRETATION VON DEUTSCHEN /r/-ALLOPHONEN

HORST ULBRICH*

1. Untersuchungsmaterial und -methode: Im Rahmen umfangreicher Untersuchungen zur *r*-Realisation im Deutschen wurden u.a. 377 /r/-Allophone¹ vom Vf. (Untersuchungsleiter) und 6 weiteren phonetisch geschulten Abhörern auditiv perzipiert und interpretiert („Hörer-Phonetik“).² Diese 377 /r/-Allophone wurden 1964 von 4 Nachrichtensprechern von Berliner Rundfunksendern im kontextuellen Zusammenhang in Wetterberichten gesprochen und für die vorliegende Untersuchung auf einem Magnettonband festgehalten. — Bei der auditiven Interpretation kam es nicht darauf an, nur eine Binärentscheidung „realisiert/nicht realisiert“ zu treffen, sondern die Realisationsqualität entsprechend dem Gehöreindruck zu bestimmen. Die auditive Perzeption wurde durch Einbeziehung eines Segmentiergerätes unterstützt. Durch beliebig wiederholbares Abhören mit Hilfe dieses Gerätes ist es möglich, weitgehend Eindeutigkeit bei der Bestimmung auch der Sprachlaute zu erzielen, die innerhalb des natürlich gesprochenen Redeflusses realisiert werden. Die klassische auditive Methode der Phonetik wird durch die Zuhilfenahme des Segmentiergerätes zur instrumentalphonetisch-auditiven. Für die instrumentalphonetisch-auditive Lautbestimmung ist die Fertigkeit des funktionellen Hörens — das ist die Fähigkeit, von der akustisch-gennematischen Perzeption auf die artikulatorisch-genetische Realisation gesprochener Laute, Wörter usw. schließen zu können — wesentliche Voraussetzung. Der eigentlichen *r*-Untersuchung ging darum ein spezielles intensives Hörtraining voraus. — Für die auditive Bestimmung wurden folgende Interpretationsqualitäten unterschieden: 1. *r*-Vibrationslaute (/r/ → [r, r̥, ʀ, R']), 2. *r*-Frikativlaute (/r/ → [ʁ] und [ʁ̥]; daneben auch [x] und — im vorliegenden Material sehr vereinzelt vorkommend — [ʁ̥]), 3. *r*-Vokallaute (/r/ → [ɐ]),

* Aus dem Institut für Phonetik und Kommunikationswissenschaft der Humboldt-Universität zu Berlin. — Direktor: Prof. Dr. G. F. Meier.

¹ // phonemische, [] phonetische Transkription.

² Vgl. A. V. Isačenko, Sprachwissenschaft und Akustik, Berlin 1966; P. Janota u. M. Romportl, Ein Beitrag zur phonetischen Methodik — Die Gehöranalyse als Bestandteil der experimentalphonetischen Forschung, *Acta Univ. Carolinae*, 1959, *Philol.* 1, S. 17—28 sowie die ausführliche Darstellung der vorliegenden Problematik und weitere Literaturangaben in H. Ulbrich, Zur *r*-Realisation im Deutschen, untersucht an der Aussprache von Rundfunksprechern und Schauspielern, Phil. Diss., Humboldt-Univ., Berlin 1966.

Martens:

Die Unterschiede zwischen 3 („vokalisch-zentral-realisiertes /r/“) gegen 4 („mit Vokal e monophthongisch realisiertes /r/“) müssen genauer formuliert werden, denn „monophthongisch mit e“ ist zu unklar. Nach den gegebenen Beispielen konnte man den Eindruck gewinnen, daß es sich auch hier um einen niedrigen Zentralvokal handelt. Bei dem Beispiel „Verkäufer“ wäre die 1. Silbe — in der gegebenen Realisierung — als Kurzdiphthong (oder Pseudodiphthong) zu werten und somit unter Gruppe 3 einzureihen.

Ulbrich:

Ad Heike: Es ging in meinem Beitrag nicht um die „Aufwertung“ der auditiven Beschreibungsmethode. Ihr Wert ist im allgemeinen anerkannt (Kongreßthema). Hinsichtlich der Bedeutung und der Notwendigkeit der erforderlichen verschiedenen Beschreibungsaspekte phonetischer Untersuchungen stehe ich im übrigen auf demselben Standpunkt wie ihn Herr Heike in seiner Arbeit „Zur Phonologie der Stadtkölner Mundart“ (1964) vertritt. Als Beleg dafür, daß in meiner Untersuchung z. B. auch der elektroakustische Aspekt berücksichtigt wurde, dienen die gezeigten Sonagramme.

Zur Frage der Übereinstimmung der Abhörergebnisse ist nochmals festzustellen, daß es eben um Hörergebnisse von Hörern (Beurteilern) mit weitgehend gleichen Voraussetzungen — also gut geschulten Hörern — ging. An Hand dieser Resultate wurden die Übereinstimmungen und Abweichungen der Urteile registriert und verglichen.

Ad Martens: Als „/er/ monophthongisch“ wird die Phonemfolge Vokal *e* und Konsonant *r* bezeichnet, wenn sie monophthongisch realisiert wird und sich gegenüber dem vokalisiertem *r* durch eine größere Zeitfunktion (über 80 ms) unterscheidet.

К ВОПРОСУ О РЕДУКЦИИ ГЛАСНЫХ В СОВРЕМЕННОМ ЛИТОВСКОМ ЯЗЫКЕ

Я. УРБЕЛЕНЕ*

Литовские безударные гласные экспериментально исследованы только на слух (Р. Эблом, В. Вайткевмчюте). На основе этих исследований говорится лишь о количественной редукции в литовском языке, в то время как качественная редукция на основе слухового восприятия отрицается. Тем не менее, судя по строю звуковой системы литовского языка безударность, очевидно, должна в какой то мере сказываться в спектральной характеристике безударных гласных.

Наше исследование, посвященное спектральному анализу долгих гласных литовского языка [a:], [æ:], [e:], [i:], [u:], [o:], дало нам возможность вскрыть их общие и индивидуальные характеристики в ударном, предударном и заударном слогах.

Результаты исследования показали, что вышеуказанные гласные в предударном и заударном слогах не утрачивают различительных признаков (компактный—диффузный, низкотональный—высокотональный, бемольный—простой), по которым они противопоставлены в системе литовского языка.

С артикуляционной точки зрения как в предударном и заударном слоге, так и в ударном, имеет место противопоставление данных гласных по переднему—заднему, высокому—низкому образованию, по огубленности—неогубленности.

Благодаря этим общим акустическим и артикуляционным характеристикам ударных и безударных гласных обеспечивается их адекватное восприятие.

Однако в спектрах долгих безударных гласных имеются особенности, отличающие их от спектров соответствующих долгих ударных гласных. Эти особенности сводятся к следующему:

1. В спектрах предударных и заударных гласных обнаруживаются сдвиги F_1 в сторону более низких, а иногда и также сдвиг F_2 в сторону более высоких частот. Следовательно, в предударном и еще в большей мере

* ЛЭФПР ВГПИ, Вильнюс.

в заударном слоге компактные гласные становятся менее компактными; диффузные — менее диффузными; низкотоновые — менее низкотоновыми, а высокотоновые — более высокотоновыми; бемольные — менее бемольными.

2. Соотнесение акустических характеристик с артикуляционными показывает, что предударные и особенно заударные долгие открытые гласные становятся менее открытыми, чем соответствующие ударные, закрытые — менее закрытыми, огубленные — менее огубленными; передние гласные произносятся с менее передней артикуляцией, а задние — с менее задней, чем в ударном слоге.

3. В спектрах предударных и заударных гласных появляются дополнительные составляющие на низких (350 гц, 500 гц, 655 гц) и реже на более высоких частотах (1150 гц, 1280 гц, 1925 гц).

4. Амплитудные взаимоотношения между максимумами спектров безударных гласных в большинстве случаев меняются, т. к. в качестве самого интенсивного максимума спектра чаще всего выступает не тот самый максимум, что и в ударном слоге.

5. Безударные долгие гласные характеризуются меньшим уровнем произносительной энергии. Произносительная энергия предударных гласных составляет приблизительно 70 %, а заударных — 30 % от общей произносительной энергии ударных гласных.

6. В безударном слоге заметно значительное сокращение длительности гласных. По сравнению с ударными длительность гласных в предударном слоге составляет, в среднем, около 70 %, а в заударном слоге — примерно 50 %.

7. В отличие от ударных долгих гласных предударные гласные менее дифтонгизированы, а в заударных гласных дифтонгизация полностью отсутствует.

8. В безударном слоге меняется характер движения основного тона. Если в ударном слоге движение основного тона может быть и восходящим и нисходящим, то в предударном слоге он чаще всего восходящий, а в заударном слоге нисходящий, причем частотный диапазон при произнесении безударных гласных уже по сравнению с ударными.

Итак, безударные долгие гласные имеют ряд особенностей, благодаря которым они отличаются от ударных. Предударные гласные отличаются в свою очередь от заударных по особым присущим им спектральным характеристикам. Как показало исследование, безударным гласным свойственна редукция (качественная и количественная), причем заударным гласным свойственна большая степень редукции по сравнению с предударными.

Редукция безударных гласных имеет в качестве своих акустических коррелятов сдвиги в местоположении максимумов, понижение уровня произносительной энергии, уменьшение длительности, изменения в движе-

нии основного тона и в частотном диапазоне. Артикуляционными характеристиками редукции служат сдвиги передней и соответственно задней артикуляции к более нейтральному положению языка; меньшее поднятие языка к небу и меньшая огубленность.

Следует, однако, подчеркнуть, что редукция безударных гласных, проявляющаяся на акустическом и артикуляционном уровне, не находит своего отражения на фонологическом уровне. Характерной чертой фонологической системы литовского языка является то, что в предударном и заударном слогах долгие гласные несут ту же функциональную нагрузку, что и в ударном слоге, т. е. система оппозиций во всех позициях остается без изменений. Следовательно, противопоставление по дифференциальным признакам гласных в трех исследованных позициях обнаруживается как на акустическом и артикуляционном, так и на фонологическом уровне. Вскрытые же в ходе акустического анализа признаки, характеризующие безударные гласные и отличающие их от соответствующих ударных гласных, являются для фонологической системы литовского языка избыточными.

SOME REMARKS ON "JUNCTURE" IN PHONOLOGICAL ANALYSIS

JOSEF VACHEK

Almost 40 years ago the well-known English phonetician Daniel Jones surprised the phonetic and linguistic world with his declaration that he could achieve "the classing of sounds into phonemes" on the basis of what was to be called later distributional grounds, i.e. without paying the slightest attention to the meaning of the texts serving as materials of this purely formal analysis. This declaration especially irritated those linguists who had been studying language as a functional system, serving primarily the communicative needs of the linguistic community. In our paper published in 1932 we believe to have demonstrated that Daniel Jones' claim cannot be justified. (Incidentally, in 1943 D. Jones' approach was also criticized by L. Hjelmslev who approved of the emphasis laid by the Prague group on the distinctive function of the phoneme.) We pointed out that if Jones' classing of sounds into phonemes were to be adequate one would have to consider at least the limits of words and morphemes within the examined contexts—and the positions of these limits are undoubtedly determined by grammatical factors, based on the facts of meaning. (The ignoring of such limits would make it impossible to solve the problem of the mono- or biphonemic value of affricate sounds in many languages.) It may be of interest that last year, in his Vienna lecture, W. Haas arrived at a conclusion analogous to ours. In answering the question of how much grammatical information is necessary for an adequate phonemic analysis he insists on the necessity of a certain minimum of grammatical prerequisites for this purpose, and as such minimum he specifies precisely the knowledge of the placement of word and morphemic limits.

While Jones' asemantic bias was due only to his approach of a practical transcriptionist of language, not to deeper considerations of linguistic theory, some 12 year later an outstanding linguistic current was to take up an approach analogous to his. This time, however, the approach was based on a thoroughly considered linguistic theory. This was the approach of the classical Yale group, especially of G. L. Trager and Bernard Bloch. Admittedly, the Yale people believed that all problems of phonemic interpretation could be solved on distributional basis alone, without any recourse to facts of meaning. Their aim, of course, was much higher than Jones's, and they well knew the pitfalls he had been unable to avoid. They were faced with the necessity of reducing the differences like ModE *nitrate*—*night rate*, ModCzech *počít* 'begin' — *podšít* 'sew under' to factors other than word or morphemic

limits, because an explanation in such terms would have implied an admission of some importance of the grammatical (and thus also semantic) factor in phonological analysis. Such an admission being ruled out in the Yale group as a matter of principle, the only course left was the postulation of a formal element existing in those positions in which the semantically orientated analysis ascertains the presence of word or morphemic limits. As is commonly known, such a postulated element indeed came to be established and termed "the juncture phoneme" (it first emerged, to our knowledge, in 1941; some other scholars have used different terms, such as "disjuncture" or, more recently, "transition").

The term has become fairly established, especially in American linguistics. It is now used more or less mechanically about the same sort of facts which Trubetzkoy, in the early 'thirties, used to call "Grenzsignale" (frontier signals), and there might be no objection to this use, if the term juncture were redefined with reference to factors of meaning with which it is obviously closely connected. One thing, however, can be regarded as certain: the qualification of junctures as phonemes will have to be dismissed. Admittedly, all allophones by which a certain phoneme becomes implemented should be mutually related, both physiologically and acoustically. Or, to put the thing in terms of Jakobson's and Halle's conception, if a phoneme constitutes a bundle of distinctive features, then the presence of all such features should be ascertainable in all variants by which that phoneme becomes implemented in concrete contexts. This conception of the phoneme has proved highly workable in most instances, but in cases of junctures it is bound to fail: an attempt to find the common features of all phonically so variegated signals of word and morphemic limits proves to be a hopeless task. An example will prove this.

The ModCzech phoneme [r] has two allophones, the syllabic and the non-syllabic one; their distribution is governed by rules also including the position of word limits. The syllabic [r], that is, is found if preceded, within the limits of the word, by a consonant and if not followed, again within those limits, by a vowel. An analyst ignoring word limits would have to establish two separate ModCz phonemes, /r/ and /r/, on the basis of oppositions like [bratr asi] 'brother perhaps' — [bratra si] 'brother to him/her', [ta krvavá smečka] 'this bloody gang' — [tak rvava: smečka] 'this fight-happy gang', etc. An orthodox distributionalist will, of course, postulate the presence of the so-called open juncture, and transcribe (1) [bratr#asi], (2) [bratra#si], (3) [ta#krvava:], (4) [tak#rvava:], etc. He will have, of course, to specify the distinctive features present in the implementations of /#/ in all the four sequences. This however, appears impossible, because /#/ in (1) is implemented by the syllabicity of /r/ preceding it, in (4), on the contrary, by the non-syllabicity of the [r] following it, and in (2) and (3) no specific phonic fact implementing /#/ can be ascertained, the implementation being the same as if no /#/ were present in the two sequences at all. Thus the "common denominator", giving the supposed juncture phoneme any phonological sense, clearly cannot be established.

In addition to this, the task would be even more hopeless if one were to interpret

those numerous sequences in which the juncture has other implementations again, such as the presence of long consonantal quantity (as in ModCz *pod domem* 'under the house' [pod:omem]), or the articulatory *détente* between a stop and a sibilant (e.g., in ModCz *pod sebou* [pot-sebou] 'under oneself'), etc. etc.—The only possible solution one can think of would be the implementation of /#/ by a potential pause. But this solution is unacceptable again: such a pause can only be phonemically characterized by the absence of any distinctive features (or, by zero features, if one prefers that term). This, naturally, stands in striking contrast to all the distinctive features otherwise employed by the Harvardian theory.

On the other hand, this zero character of the supposed juncture phoneme is in full conformity with the classical Prague conception of Trubetzkoy's frontier signals. The fact that such a potential pause can never occur inside a morpheme but only between words (and, though less frequently, between morphemes constituting a word) repeatedly endorses the importance of the factor of meaning in phonological analysis, the factor whose disregard actually gave rise to the concept of "juncture phoneme". This concept is clearly as unjustified as the deliberate refusal to take into consideration in phonological analysis what is in fact the *raison d'être* of language, i.e. its communicative function, consisting in the reference by phonic means to extralingual reality, reference based on the existence of the meaning of words and morphemes.

DISCUSSION

Singh:

In searching for the acoustic correlates to *sandhi* I have found that almost 90% of the *sandhi* rules given in Sanskrit grammars have almost one-to-one acoustic correlates. For example, the word compound *gri:f* is a combination of two independent words *gri* + *i:f*. By using electronic gates when the word *gri:f* was truncated at the juncture it was found by using the psychophysical method of adjustment that two independently said words *gri* and *i:f* were perceived by the listeners not different from the truncated words *gri* and *i:f*.

Vachek:

I gladly admit the delimitative function of "junctures". But exactly this function is closely associated with meaning (grammatical or lexical); thus this latest conception is free from the drawbacks of the old one I criticized here. It should be added that it is especially the mechanical use of the term "juncture phoneme" by non-Yale scholars which is open to criticism, as such use is in no way consistent with the classical European, or the Harvardian, conceptions of phoneme and phonemics.

NASALIZATION IN CREOLE FRENCH

ALBERT VALDMAN*

This paper will attempt to show that a more satisfactory analysis of the phonological function and phonetic realization of nasalization in Creole French may be obtained if one departs from a view of phonological analysis anchored on the concept of the phoneme and if one adopts one in which the primary concerns are the indication of the choices available to the native speaker at any point in a sentence and the specification in the most economical way of the phonetic realization of sentences.¹ The data examined here are drawn from Haitian Creole French (HC) but the solution proposed applies on the whole *mutatis mutandis* to all other Creole French dialects.

The most explicit analysis of HC (Hall, 1953) posits five nasal vowels [*ĩ ù ě õ ã*], all of which except for [*ã*] bear a one-to-one relationship to some oral vowel counterpart. Seven oral vowels are assumed for rural varieties of HC including a tense versus lax contrast for the mid vowels, that is:

Oral Vowels			Nasal Vowels	
<i>i</i>	(<i>ü</i>)	<i>u</i>	<i>ĩ</i>	<i>ũ</i>
<i>é</i>	(<i>æ</i>)	<i>ó</i>		
<i>è</i>	(<i>ê</i>)	<i>ò</i>	<i>ẽ</i>	<i>õ</i>
—	<i>a</i>		<i>ã</i> = [ə]	

Many dialects exhibit three front rounded vowels ([*ü œ ê*]) but these do not bear on the analysis of nasalization and will be left out of consideration.

HC nasal vowels [*ĩ*] occur in more environments than do those of Standard French; in particular, they occur widely before syllable- and word-final nasal consonants (N). It might be supposed therefore that in HC nasalization assumes a greater differentiative role in the vowel system, but in fact the reverse obtains for the following reasons.

First, the high vowels [*i*] and [*ũ*] occur infrequently and, except for the first person plural inclusive pronoun *nou* and the indefinite determiner *youn* as well as words pertaining to the vaudoun rites (*oungan, ounfò*), they occur only before N,

* Department of Linguistics, Indiana University.

¹ The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

e.g. *bounda* 'buttock', *pinga-ou* 'watch out!', *kachimbo* 'pipe'.² Second, while in HC, unlike Standard French, oral and nasal vowels co-occur before N, the number of attested contrasts is small. Table 1 shows a representative set of such contrasts and also illustrates the distribution of the mid vowels, both oral and nasal.

Chart 1. Distribution of Oral and Nasal Low and Mid Vowels

CV	C \tilde{V}	CVN	C \tilde{V} N	CVC	C \tilde{V} C
<i>šē</i> 'dear'	<i>šē</i> 'dog'	<i>šēn</i> 'chain' ~ <i>žēn</i> 'young'	<i>šēn</i> 'chain' <i>žēn</i> 'fast'	<i>pēs</i> 'pest' <i>šat</i> 'cat' <i>pat</i> 'paw'	<i>pēs</i> 'pliers' <i>pāt</i> 'slope'
<i>sa</i> 'that'	<i>šā</i> 'field'	<i>šam</i> 'charm' <i>pan</i> 'breakdown'	<i>šām</i> 'room' <i>pān</i> 'hang' <i>vān</i> 'sell'		
<i>pó</i> 'pot' <i>pò</i> 'harbor'	<i>vā</i> 'wind' <i>pō</i> 'bridge'	<i>pōn</i> 'lay eggs' <i>bōn</i> 'boundary' <i>bōn</i> 'maid' — <i>mōn</i> 'hill'	<i>pōn</i> 'lay eggs' <i>bōn</i> 'good' <i>bōn</i> 'maid' <i>mōn</i> 'world'		
<i>mó</i> 'word' <i>mò</i> 'dead'				<i>mòd</i> 'fashion'	<i>mōt</i> 'watch'

Third, one of the striking characteristics of HC is the diffusion of nasalization over contiguous sequences of vowels and consonants. This results in three types of alternations described variously as neutralization, free variation, and allophonic or phonemic effects of juncture or disjuncture (Hall, 1950; 1953; Taylor, 1947): (1) stops and fricatives are pre-nasalized optionally after nasal vowels, e.g. [*š^mpil*] 'a lot', [*š^{ndyē}*] 'Indian' [*š^{ngē}*] 'vaudoun priest'; (2) vowels are partially or fully nasalized preceding or following N, e.g. [*tuné*][*tūné*][*tūnē*] 'to turn'; (3) word-final \tilde{V} may be followed by velar closure, e.g. [*lō*][*lō^v*] 'long'.

For any phonological analysis of HC to attain at least observational adequacy, it must account for contrasts of the type [*bōn*] 'good' versus [*bōn*] 'boundary', and it must also indicate where nasalization is optional, obligatory or not permissible. This paper proposes to meet these requirements by assigning to each HC grammatical form a single abstract underlying representation consisting of a string of distinctive feature matrices and formulating a set of ordered rules which transform the underlying representations into observable phonic features. The proposed analysis differs

² To represent HC forms we employ a modified version of the Faublas-Pressoir (ONEC) orthography. The following typographical conventions will be observed: forms cited in the orthography will appear in italics; forms between // are the result of the application of one or more PR; English glosses appear between ' '; [] indicate phonetic transcription; underlying representations appear in roman.

from existing ones in that nasalization is treated as a feature absent from the underlying representation and determined by length. The following underlying vowels are posited for HC:

	i	u	e	o	a	á	é	ò
diffuse	+	+	—	—	+	+	—	+
compact	—	+	—	+	—	—	—	+
grave	—	—	—	—	—	+	+	+
long	—	—	—	—	—	—	—	—

To derive the nasal vowels we apply two phonological rules (PR). PR1

$$\left[\begin{array}{c} V \\ \text{—long} \end{array} \right] \rightarrow [+ \text{nasal}] \text{ in the env. } - \left\{ \begin{array}{l} \text{NX\#} \\ \text{NCX} \end{array} \right\}$$

(where N means any nasal consonant, C is any consonant, X any sequence of vowels or consonants or null and # word boundary) transforms any—long vowel followed by a nasal consonant to a \tilde{V} , e.g. *šanm* 'room' \rightarrow [*šānm*], *šēn* 'chain' \rightarrow [*šēnn*], *šēn* 'dog' \rightarrow [*šēn*]. PR1 of course does not apply to + long vowels which remain non-nasal before N, e.g. *bōn* 'boundary' \rightarrow [*bōn*]. PR2, which can only be applied once to any grammatical form, deletes Ns that occur after a \tilde{V} and before #

$$N \rightarrow \emptyset \text{ in the env. } \left[\begin{array}{c} V \\ + \text{nasal} \end{array} \right] - \left\{ \begin{array}{l} (N) \\ (C) \end{array} \right\} \#$$

(where C means any consonant), e.g. *šēnn* # \rightarrow [*šēn*], *šānm* # \rightarrow [*šām*], *šēn* # \rightarrow [*šē*] but *telefōne* # 'to telephone' \rightarrow [*telefōne*].

Two types of \tilde{V} emerge from this analysis which we label *primary* and *secondary*. The former are derived from underlying sequences VN and occur before non-nasal consonants or before # while the latter are derived from underlying sequences VNN and occur before N. This distinction is reflected at the phonetic level where, according to preliminary spectographic data, secondary \tilde{V} are not as fully or clearly nasalized as primary \tilde{V} ; as has also been shown secondary \tilde{V} are in some instances replaced by oral vowels. It is tempting to equate the distinction primary versus secondary \tilde{V} to the distinction \tilde{V} by cancellation (*annulation*) versus \tilde{V} by damping (*amortissement*) respectively proposed by Pierre Delattre (1967). For Delattre maximally distinctive \tilde{V} s result from the reduction of the strength of the first formant and requires in addition to the lowering of the velum a pharyngeal cavity of equal size to the velic cavity formed above the velum. This condition is met by \tilde{V} with a first formant of about 600 cps, the low mid vowels. Thus three interesting generalizations emerge from the proposed analysis: (1) only the mid vowels [*ē* *ò*] and the low vowel [*ā*] are primary \tilde{V} ; (2) the feature + long which blocks nasalization occurs in many words which ma

be considered peripheral to the core HC lexicon; (3) the presence of velar closure after final \tilde{V} and of pre-nasalization of stops and fricatives after \tilde{V} may be accounted for by the non-application of PR2, which allows the N to appear at the phonetic level.

The assignment of distinctive feature status to length rather than to nasalization also accounts for the disruption of complementary distribution between high-mid (tense) and low-mid (lax) vowels, that is, that together with [m \acute{o}] 'word' and [m \grave{o} d] 'fashion' we find [m \grave{o}] 'dead'. PR3—5 displayed below assign + or — diffuse values (in articulatory terms, tense versus lax) to vowels which are unaffected by PR1—2

$$PR3: \begin{bmatrix} - \text{compact} \\ + \text{long} \end{bmatrix} \rightarrow \begin{bmatrix} - \text{diffuse} \\ - \text{compact} \\ (+ \text{long}) \end{bmatrix}$$

$$PR4: \begin{bmatrix} + \text{compact} \\ + \text{long} \end{bmatrix} \rightarrow \begin{bmatrix} + \text{compact} \\ (+ \text{long}) \end{bmatrix}$$

$$PR5: \begin{bmatrix} - \text{compact} \\ - \text{long} \end{bmatrix} \rightarrow \left\{ \begin{array}{l} \begin{bmatrix} - \text{diffuse} \\ - \text{long} \end{bmatrix} \\ \begin{bmatrix} + \text{diffuse} \\ - \text{long} \end{bmatrix} \end{array} \right. \text{ in the env. } \left. \begin{array}{l} \text{--- } C_1(CV)_0 \# \\ \text{--- } (CV)_0 \# \end{array} \right\}$$

(where C_1 means one and only one consonant and $(CV)_0$ one or more sequences of vowel plus consonant).

All—compact vowels become—diffuse (lax) whereas short vowels become—diffuse (lax) or + diffuse (tense) depending on whether they occur in a checked or free syllable respectively. As a result tense and lax vowels contrast in open syllables: mo [m \acute{o}] 'word' versus m \acute{o} [m \acute{o}] 'dead' or vole [v \acute{o} lé] 'to fly' or 'to steal' versus volé [v \acute{o} lé] 'thief' but mod [m \acute{o} d] 'fashion' and pes [p \acute{e} s] 'pest'. The + compact long vowel \acute{a} remains unchanged, except that, as for the other underlying long vowels, it is realized with optional length and, in some dialects, with a following [r].

REFERENCES

- Delattre, P. Divergences entre nasalités vocalique et consonantique (forthcoming).
 Hall, R. A. Jr. *Haitian Creole: Grammar, Texts, Vocabulary* (Memoir No. 43 of the American Folklore Society and Memoir No. 74 of the American Anthropologist), Philadelphia, 1953.
 ———. Nasalization in Creole French, *Modern Language Notes* 65 (1950) 474—478.
 Taylor, D. Phonemes of Caribbean Creole, *Word* 3 (1947) 173—179.

DISCUSSION

Annan:

In my opinion it might be more practical to use a non-generative type of analysis but rather to employ a prosodic approach as practised in the London school which would allow the problem to be seen in its position in phonology.

Carton:

Je signale que dans certains patois et en français populaire dans le Nord de la France, le nombre des nasales du système est réduit à deux: un \tilde{a} teinté de \acute{o} et un \tilde{e} teinté de \acute{e} , allant ainsi jusqu'au bout des possibilités de la langue.

Goodman:

Why set up a feature of length for the vowel a when it is not present anywhere in the language and serves merely as a device to block nasalisation in such forms as $\acute{s}am$ "charm" versus $\acute{s}\tilde{a}m$ "room", a phenomenon just as well accounted for by the nasal versus oral vowel contrast.

LES PERSPECTIVES DE LA LOGOGRAPHIE

B. VALLANCIEN—A. DJOURNO—A. MARTIN—B. FAULHABER

DEFINITION

C'est une méthode qui permet d'obtenir de toute voyelle un diagramme caractéristique indépendant à la fois de la hauteur du son émis et du timbre de la voix qui l'émet.

Les formants d'un son vocalique sont constitués de bandes de fréquences qui définissent ce son par opposition aux autres sons de voyelle.

En réalité seul un spectre continu caractérise complètement ce son, mais dans le spectre on ne retient que les fréquences ou groupes de fréquences plus intenses: „les crêtes“ qui identifient la voyelle (fig. 1). En fait, deux zones surtout sont respon-

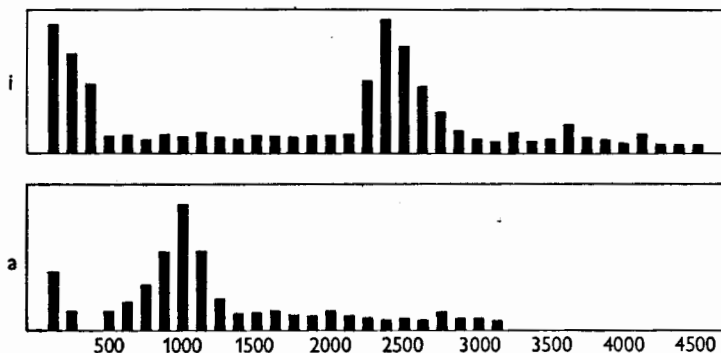


Fig. 1. Spectres de fréquence de deux voyelles. Le „i“ possède deux zones caractéristiques („formants“) dont une autour de 2500 hertz. Pour le „a“ les deux zones sont pratiquement confondues en une zone entre 700 et 1200 (type compact) (d'après Fletcher, *Speech and Hearing in Communication*, New—York, 1953).

sables du type vocalique, les autres déterminent les nuances individuelles. Il existe ainsi un champ de liberté qui justifie la possibilité d'émettre la même voyelle à différentes hauteurs. C'est pourquoi il faut bien distinguer les fréquences formantielles du son fondamental ou fréquence de récurrence.

Le son fondamental est déterminé par le nombre des occlusions glottiques à chaque seconde. Les formants le sont par l'excitation des résonateurs pharyngiens, buccaux et nasaux mis en service par le son laryngé. Les formants sont, en fait, des harmoni-

ques du fondamental laryngé et, lorsque celui-ci varie, leur fréquence varie aussi mais seulement dans les limites de la bande formantielle faute de quoi la voyelle perdrait son intelligibilité.

Pour reconnaître une voyelle il faut et il suffit que ses formants soient situés dans leurs zones de fréquence de prédilection et avec une intensité relative convenable.

On peut analyser sommairement une voyelle et rendre visibles ses deux formants en les extrayant du son émis à l'aide de deux filtres acoustiques que l'on fait agir selon deux axes rectangulaires sur la tache lumineuse de l'écran oscillographique.

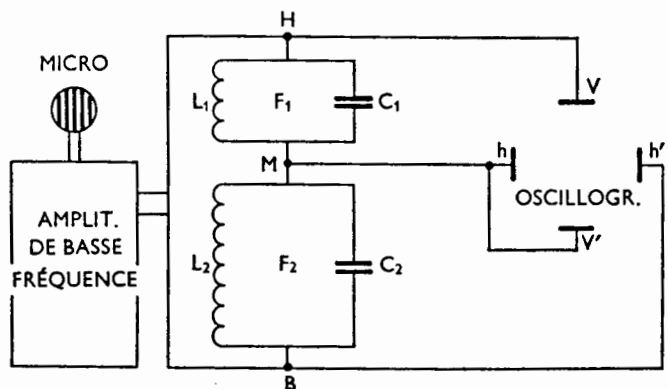


Fig. 2. Schéma de montage d'un logographe. F 1: filtre développant des tensions notables dans les zones de fréquence des sons aigus (formant haut). F 2: filtre „bas“, provoquant un balayage horizontal (formant bas). L 1—L 2: bobines à noyau de fer; C 1—C 2: Condensateurs au papier.

Le matériel est simple: un microphone, un amplificateur, un oscillographe et un double filtre qu'on peut construire soi-même. Deux petits condensateurs — et deux bobines de self — constituent les deux filtres qui doivent présenter une fréquence de résonance et un amortissement convenables.

Pour une bonne discrimination, le maximum de la courbe de résonance de chaque filtre doit correspondre respectivement à la limite inférieure de la zone du formant bas et à la limite supérieure de la zone du formant haut (fig. 2).

Pour le formant bas, nous avons utilisé une bobine à noyau de ferocube de 1 henry environ de 110 ohms de résistance et un condensateur papier de 0,10 microfarad.

Pour le formant haut: bobine à noyau de ferocube de 70 millihenrys environ et de 22 ohms de résistance et condensateur papier de 0,05 microfarad.

Fig. 3 — Le filtre F1 correspondant au formant haut provoque une déflexion verticale du spot lorsqu'il développe à ses bornes une tension électrique.

Les basses fréquences de l'autre formant traversent assez facilement F1 mais développent au contraire une différence de potentiel notable aux deux bornes du filtre F2 c'est-à-dire entre B et M. Elles produisent un balayage horizontal du spot. Ainsi ces deux différences de potentiel alternatives vont se traduire par une figure

lumineuse qui sera une courbe fermée chaque fois que le bruit sera périodique, comme dans un son de voyelle.

Cette figure sera décrite par le spot plus ou moins rapidement suivant la fréquence du fondamental mais sa forme ne dépendra que des phases relatives des constituants du son complexe analysé c'est-à-dire des formants qui caractérisent la voyelle. On l'appellera un logogramme.

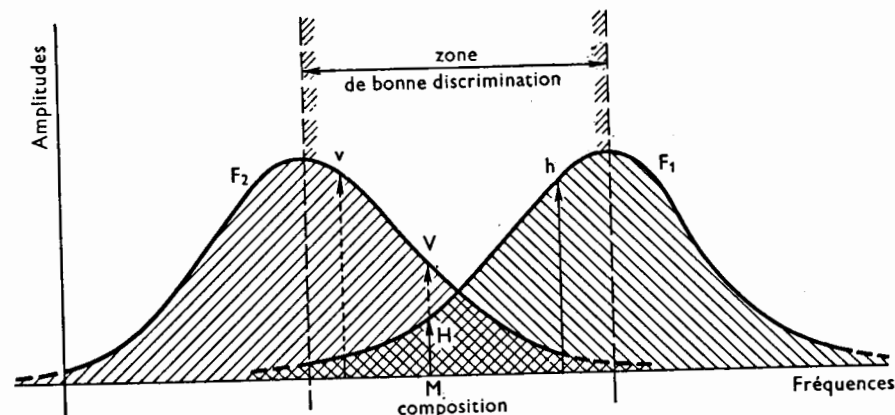


Fig. 3. F, 1: courbe de résonance du filtre haut, F, 2: celle du filtre bas, v: amplitude de la déviation horizontale (pour un son grave), h: amplitude de la déviation verticale (pour un son aigu), MH: amplitude de la déviation horizontale et MN amplitude de la déviation verticale pour un son de hauteur moyenne.

1°) — CAS DE SONS PURS

Pour un son pur dit sinusoïdal si la fréquence est très aiguë le logogramme est une ligne verticale.

Pour un son de fréquence très grave le logogramme est une ligne horizontale.

2°) CAS DE SONS COMPLEXES

C'est-à-dire constitués par des fréquences différentes.

A — Rapport des fréquences

Si une fréquence est le double de l'autre elle aura pour effet:

- de redresser la courbe puisqu'elle correspond à un son plus aigu.
- de lui donner un aspect plus sinueux, dentelé puisque le spot est sollicité deux fois plus souvent dans un sens que dans l'autre.

B — Déphasage

Si le spot est sollicité verticalement et horizontalement par un son de même fréquence mais décalé dans le temps il ne décrira plus une droite mais une ellipse

plus ou moins allongée qui peut devenir un cercle si les deux amplitudes sont égales et le déphasage d'un quart de période.

Dans les logogrammes de voyelle deux harmoniques déphasés engendreraient des courbes plus ou moins compliquées. Or, cette complexité est très précieuse puisqu'elle va permettre de caractériser la voyelle (fig. 4).

On peut même obtenir des images des semi-voyelles et reconnaître à la limite des consonnes grâce aux figures de transition qui, par voisinage, modifient la configuration de la voyelle.

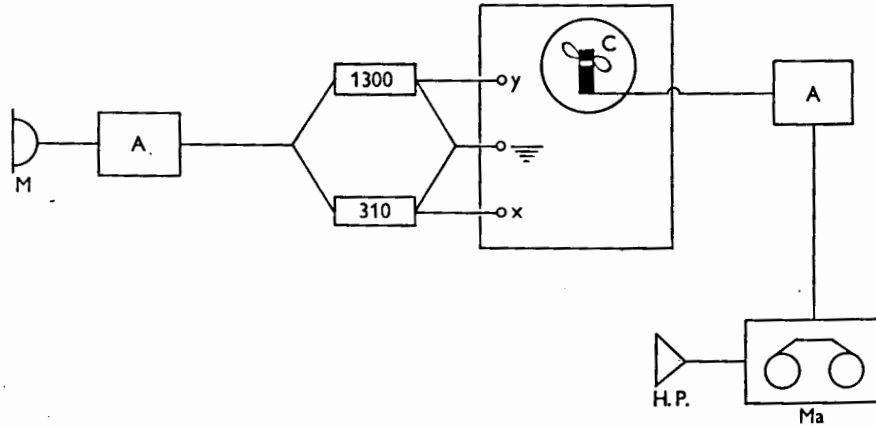


Fig. 5. Dispositif permettant la synthèse acoustique de signaux à l'aide d'une cellule photo transistor et d'un magnétophone. M = Microphone, A = Amplificateur. 310, 1300 = fréquences nominales des filtres. X et Y = entrées de l'oscillographe. Ma = Magnétophone, Hp = Haut-parleur, C = cellule.

Malgré tout, une grande complexité peut nuire à la lecture du tracé de la voyelle et empêcher d'identifier la consonne qui précède ou qui suit. Aussi, nous avons cherché à mettre au point des logogrammes *dérivés* qui tiendraient compte des variations des formants en amplitude et en phase de façon à traduire l'influence de la consonne sur la voyelle. Cette étude est en cours.

De même, en vue d'un alphabet logographique nous avons entrepris de standardiser les amplitudes pour avoir des diagrammes de même dimension.

Enfin, nous nous efforçons d'apurer le signal en supprimant les amplitudes parasites, restes de formants, inutiles et partiellement transmis par les filtres.

Plus intéressante nous a paru la vérification du bien fondé de notre analyse en effectuant une synthèse acoustique des signaux. La solution est très simple: une cellule photo-transistor placée en face de la trace oscillographique injecte sa tension aux bornes d'un magnétophone. Le résultat prouve que tout ce qui est utile à l'intelligibilité de la parole figure dans ce signal. Ajoutons que ce n'est pas le cas si les fréquences nominales des filtres sont choisies arbitrairement et si les conditions expérimentales énoncées ne sont pas respectées. Audition de l'enregistrement ainsi réalisé: émission de voyelles et d'un texte.

BIBLIOGRAPHIE

- A. Djourno, Analyse oscillographique des voyelles. *Soc. de Bio.* 1959, 153 : 197.
 B. Faulhaber, B. Vallancien, H. Lenoir, Enregistrement électro-acoustique de signaux lumineux modulés par le son de la parole. Intérêt de la méthode en phoniatry. — *C. R. 26e Cong. Sté Fce Phon.* Oct. 1967.
 H. Lenoir, Image de quelques voyelles françaises obtenue par diagramme oscilloscopique. — *Rééduc. Orthop.* 1967.
 B. Malmberg, La Phonétique.-collec. Que sais-je, PUF, Paris 1954.
 B. Malmberg, Le problème du classement des sens du langage et quelques questions connexes. — Lund (Suède) 1952.

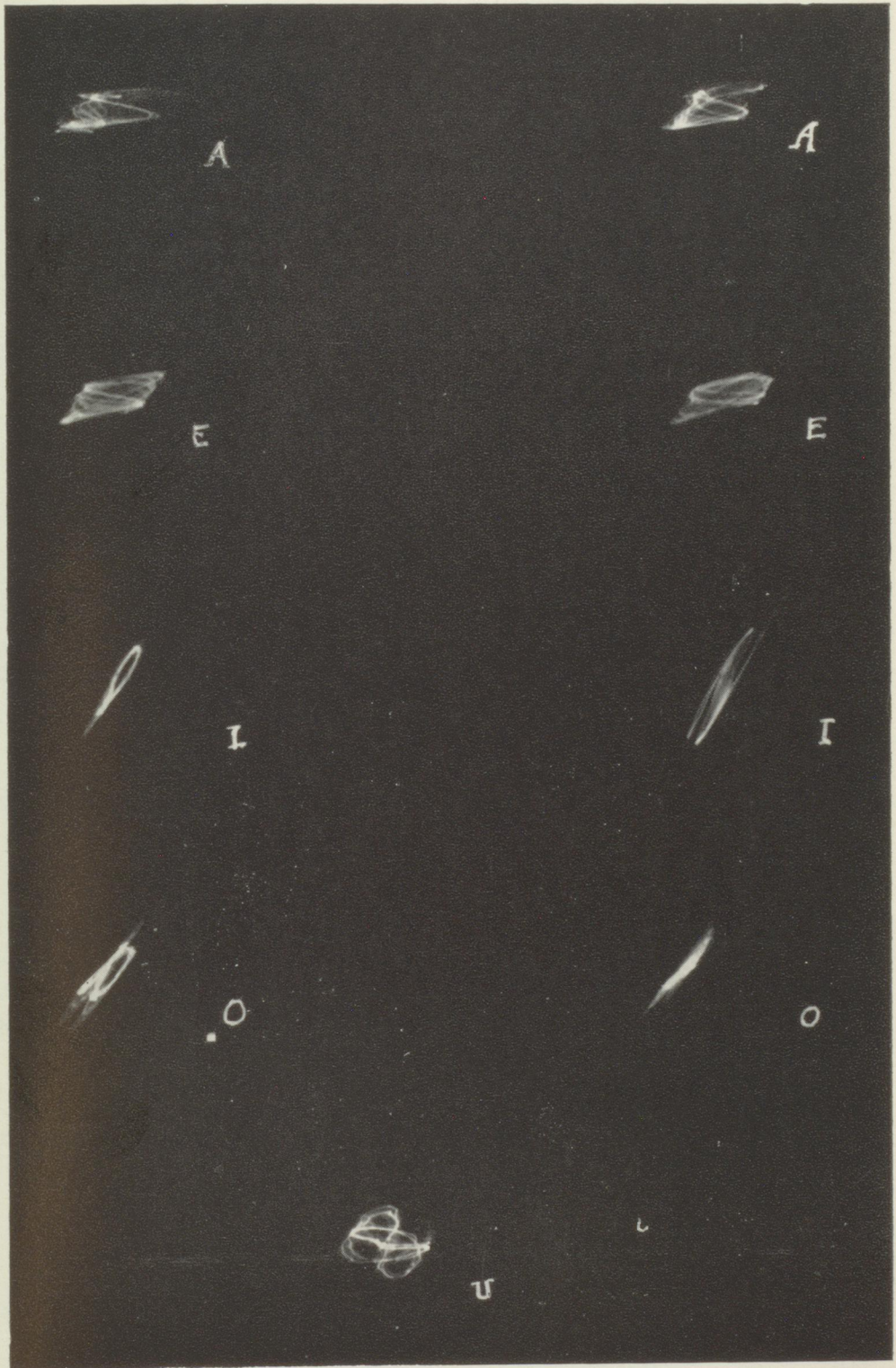


Fig. 4. Logogramme de différentes voyelles. Chaque son de voyelle a été émis par deux personnes différentes. Le logogramme permet d'identifier la voyelle énoncée.

RULES ORDERING AND EVOLUTION OF PHONEMIC SYSTEMS

E. VASILIU*

1. We shall give a brief characterisation of the consonantism of dialects spoken in the points 219, 228, 250 and 260 of ALR.

- (a) The primary dental oral stops $[t, d]$ are represented by the affricates $[\acute{c}, \acute{d}]$ (phonetic transcription of ALR) in the points 250 and 260.
- (b) The primary dental oral stops $[t, d]$ are represented by the palatal stops $[k, \acute{g}]$ in the points 219 and 228.
- (c) The primary palatal stops $[k, \acute{g}]$ are represented in 219, 228, 250, 260 by the affricates $[\acute{c}, \acute{d}]$ (phonetic notation of ALR).
- (d) The primary affricates $[\acute{c}, \acute{g}]$ are represented in 219, 228, 250, 260 by $[\acute{f}^1, \acute{z}]$ which are distinct from the primary non-palatal $[\acute{f}, \acute{z}]$.

In contradistinction to the literary dialect and the other dacoromanian dialects, the phonemic systems of the dialects above referred to are characterized by:

- (α) the absence of primary palatal stops $[k, \acute{g}]$ (points 219, 228, 250, 260) as a consequence of (c);
- (β) the absence of primary affricates $[\acute{c}, \acute{g}]$ (points 219, 228, 250, 260) as a consequence of (d);
- (γ) a secondary set of affricates $[\acute{c}_2, \acute{g}_2]^1$ (in 250, 260) as a consequence of (a);
- (δ) a secondary set of affricates $[\acute{c}_2, \acute{g}_2]$ and a secondary set of palatal stops $[k_2, \acute{g}_2]$ (points 219, 228) as a consequence of (c) and (b), respectively.

2. In terms of a transformational phonology in Halle's 1964 form, the dialects mentioned under 1. may be described by supplementing the set of rules which are common for all the dacoromanian dialects with some phonetic rules accounting for the characterization given by (a)—(d) (see Vasiliu, 1967a, b) namely

$$\begin{array}{lcl}
 (1) & [\acute{c}, \acute{g}] & \rightarrow & [\acute{f}^1, \acute{z}] & \left\{ \begin{array}{l} [e] \\ [g] \\ [\acute{t}] \\ [j] \\ [\acute{j}] \end{array} \right. \\
 (2) & [t, d] & \rightarrow & [k, \acute{g}] & \\
 (3) & [k, \acute{g}] & \rightarrow & [\acute{c}, \acute{g}] &
 \end{array}$$

* University of Bucharest.

¹ The phonetic distinction between $[\acute{c}, \acute{d}]$ and $[\acute{c}_2, \acute{g}_2]$ is not phonologically relevant and therefore we do not mention the distinction in our phonemic representation.

Rule (1) must be applied first, in order to keep the difference between the secondary affricates [\hat{c}_2, \hat{g}_2] and the primary ones.

In 250, 260, rule (2) has to be ordered *before* (3), as well as in 219, 228, rule (2) has to be ordered *after* (3).

3. We shall try now to answer the following question: how this different ordering of rules has to be historically interpreted?

In 250, 260, the palatalized [t, d] was acoustically identified with primary [k, g] so that both series fell into the domain of rule (3). That means rule (3) was *acquired* after the palatalization of [t, d].

In 219, 228, the confusion between [k, g] and [\hat{c}, \hat{g}] is earlier than the palatalization of [t, d] and then the palatalized [t, d] remain distinct from [\hat{c}, \hat{g}].

We could now establish the subsequent hypothesis:

The area represented by 250, 260, 219, 228 belonged at an earlier stage to the larger area, where [t, d] were palatalized in the form [k, g]. This first assumption is confirmed by the fact that 250, 260 make an isolated area (with [\hat{c}, \hat{g}] palatalization of [t, d]) within the area where [t, d] are consistently palatalized in the form [k, g]. Earlier than the palatalization in 219, 228, was the change (c), expressed by rule (3). In this way, 219 and 228 made a distinct sub-area (with palatalized [t, d] distinct from [\hat{c}, \hat{g}]) within 250, 260, 219 and 228.

The pronunciation [\hat{c}, \hat{g}] of every [k, g] was 'imported' by the dialect 250, 260 in a later stage from the dialect 219, 228. For in 250, 260 the speakers had the same sound perception of palatalized [t, d] and of primary [k, g], when the new pronunciation (from 219, 228) was adopted, the merger of palatalized [t, d] and the primary [k, g] into [\hat{c}, \hat{g}] was quite normal.

The case above discussed gives an example of one of the possible concrete evolutive interpretations of the purely formal concept of different ordering of the same rules.

REFERENCES

- Halle, 1964 = M. Halle, *Phonology in Generative Grammar*, in J. A. Fodor & J. J. Katz, *The Structure of Language: Readings in the Philosophy of Language*, Englewood Cliffs, N.Y. Prentice Hall, pp. 334—352.
- Vasiliu, 1967a = E. Vasiliu, *Transformational vs. Biunique Phonemic Typology*, in *Phonologie der Gegenwart*, Vorträge und Diskussionen anlässlich der Internationalen Phonologie-Tagung in Wien, 30. VIII.—3. IX. 1966, Graz—Wien—Köln, 1967, pp. 254—261.
- Vasiliu, 1967b = E. Vasiliu, *Considérations typologiques sur la phonologie transformationnelle des dialectes dacoroumains*, in *Cahiers de linguistique théorique et appliquée*, vol. IV, 1967, pp. 253—260.

АКУСТИЧЕСКИЕ КОМПОНЕНТЫ РЕЧЕВОГО СИГНАЛА И ИХ ФУНКЦИИ

В. А. ВАСИЛЬЕВ*

1. Автор настоящего сообщения понимает речевой сигнал не как единицу обычного кода типа азбуки Морзе и т. п., а как составную часть второй сигнальной системы человека, выявленной И. П. Павловым, как „сигналы первых сигналов“ действительности, как „вторые сигналы в виде слов“.

2. Язык как реализующееся в речи важнейшее средство человеческого общения и единственное орудие мысли является второсигнальной семиотической системой и не может быть отождествлен с обычным кодом, так как это означало бы отрыв языка от мышления. Всякий обычный код предполагает кодирование сообщения из языковой формы в кодовую знаками этого кода. Если же считать и язык кодом, то остается неизвестным, из какой же доязыковой формы сообщение перекодируется в языковый код.

3. Язык внешне сходен с кодом лишь в том смысле, что оба могут существовать только в материальной форме. Однако если для обычного кода абсолютно безразлична материальная форма его существования (звуковолновая, световолновая, электроимпульсная, радиоволновая), то язык в *принципе* может существовать только в звуковолновой материальной форме. Поскольку же звуковая материя языка — физическое, а именно акустическое явление, то и речевой сигнал может быть только акустическим.

4. Акустическая структура речевого сигнала является комплексной — она представляет собой неразрывное единство нескольких изменяющихся во времени компонентов, каждый из которых в действительности не существует отдельно от других, а может быть вычленен лишь условно — в целях анализа и описания с различных точек зрения, главным образом, с функциональной.

5. Первый и основной из этих компонентов можно назвать спектральным, формантным или высокочастотным. Он представляет собой акустическое соответствие того компонента звукового строя языка, который образуется фонетическими единицами, известными на артикуляционном и перцептивном уровнях как звуки речи.

* ЛЭФИПР 1 МГПИИЯ им. М. Тореца, Москва.

6. Второй компонент, образуемый частотой основного тона (ч. о. т.), т. е. низкочастотными колебаниями голосовых связок по всей их длине, предлагается называть высотным или низкочастотным. Он образует основу фразовой интонации.

7. Третий компонент — силовой. Он лежит в основе динамического ударения и также участвует в образовании интонации.

8. Четвертый компонент — временной, или темпоральный. Он проявляется в длительности звуков речи.

9. В акустической структуре речевого сигнала можно гипотетически постулировать и наличие тембрального компонента. Механизм его производства и физическая природа еще не исследованы, но он, по-видимому, образуется парциальными колебаниями голосовых связок, создающими обертоны, но иные, чем те, которые определяют качество гласных и индивидуальный тембр голоса, хотя и находятся с последними и с ч. о. т. в весьма сложных взаимоотношениях. Функция этого компонента заключается в различной эмоциональной окраске произносимых предложений.

10. У всех компонентов наблюдается функциональный изоморфизм — каждый из них выполняет одни и те же три общие функции, неразрывно связанные между собой и вычленимые только в целях анализа.

11. Их первая функция, *sine qua non*, — конститутивная. Она заключается в том, что каждая реализующаяся в речи единица языка — алломорфа, безморфемное слово, словоформа, словосочетание, предложение (фраза) — обязательно образуется этими компонентами в том или ином их сочетании. Так, все языковые единицы образуются звуками речи, каждый из которых обязательно имеет ту или иную интенсивность (силу) и ту или иную длительность, а гласные, сонанты и звонкие шумные согласные — также и ту или иную ч. о. т.

Слово и словоформа (даже односложные) обязательно имеют словесное ударение, а словосочетание и фраза — ту или иную интонацию.

12. Вторая функция, возможная только на основе первой — дистинктивная (морфемодистинктивная, словодистинктивная, формодистинктивная, фразодистинктивная). Изменения в каждом из акустических компонентов речевого сигнала могут производиться говорящим специально, намеренно, целенаправленно, а именно в дистинктивных целях — для различения языковых единиц. В результате этого каждая этическая единица становится и эмической. Так, звуки речи являются аллофонами сегментных фонем; дистинктивные изменения в длительности одного и того же звукового типа образуют слово-или формохроному (чешск. *dal* — дал и *dál* — дальше); дистинктивные изменения в силе и месте ударения образуют слово — и формоакцентемы (русск. *му́ки* — муки́, *ру́ки* — руки́), а также и фразоакцентемы („Вот моя племянница — Лю́ся“ и „Вот моя племянница, Лю́ся“); дистинктивные изменения в ч. о. т. внутри изолированного слова в тональных языках

образуют слово- и формотонемы, а внутри предложения, например, в русском — фразотонемы (утверждение „Он там“ — вопрос „Он там?“); совокупность целенаправленных изменений в нескольких компонентах интонации (высотном, силовом, темпоральном и тембральном) образует наивысшую фонологическую единицу языка — интоне́му (вопрос „Какие гербы?“ — восклицание „Какие гербы!“).

13. Третья функция — рекогнитивная. Она заключается в употреблении надлежащих „алло-“ каждой „-эмы“ в надлежащей позиции и ситуации. Это обеспечивает наиболее легкое и скорое узнавание языковых единиц слушающим, а следовательно, и беспрепятственное понимание им устной речи. Нарушение говорящим рекогнитивной функции не ведет к смещению языковых единиц, как это имеет место при нарушении дистинктивной функции, но отвлекает внимание слушающего от содержания речи на ее необычное звучание и тем самым затрудняет ее понимание.

Иногда опознавание одной фонемы обеспечивается не ее собственными дистинктивными признаками, а скорее употреблением надлежащего аллофона соседней фонемы (например, удлиненных и кратких аллофонов английских фонем перед звонкими и глухими шумными согласными, как в *cab* — *cap*). Таким образом, дистинктивная функция вообще может проявляться только через рекогнитивную.

ФОНЕМНАЯ СТРУКТУРА СЛОВА В СОВРЕМЕННОМ АНГЛИЙСКОМ И ЛАТЫШСКОМ ЯЗЫКАХ

М. ВЕЦОЗОЛ*

В данной работе исследованы характерные особенности фонемной структуры слова современного английского и латышского языков (АЯ и ЛЯ).

Под фонемной структурой слова понимается: состав и количество фонем, распределение фонем в слове и возможные сочетания фонем в различных положениях в слове.

При исследовании фонемной структуры слова обоих языков были проанализированы тексты различного характера, а также использованы разные словари (как D. Jones, An English Pronouncing Dictionary, E. Turkina, Latvian-English Dictionary и т. д.)

В докладе излагаются главные результаты этого исследования.

1. В обоих языках имеются слова, в которых *состав фонем* может быть разным, но *количество* одинаковым, например, в АЯ [mi:n] mean, [led] lead, [tɒp] top, [ri:b] rib, в ЛЯ [grab] grab, [likt] likt.

2. В обоих языках имеются слова, содержащие *один и тот же фонемный состав и одно и то же количество фонем*, однако *распределения этих фонем различны*, например, в АЯ: [aɪd] eyed, [daɪ] die, [da:k] dark, [ka:d] card; в ЛЯ: [apɪur] apɪur, [putra] putra.

3. В обоих языках слова *могут состоять из одной фонемы*, т. е. могут иметь *самое меньшее число фонем* в слове, напр., в АЯ [ɜ:] awe, or, ore, [eə] air, [aɪ] I;

в ЛЯ имеется только несколько однофонемных слова, напр., [èi] ei, [ài] ai! [aù] au (u).

Самое большое количество фонем в слове в АЯ — 19, напр., [ə'kɒmpəni'ments'bilɪti] accompanimentability, а в ЛЯ — 23, напр. ['struktura:l'indiferentais] strukturalindiferentais.

4. В обоих языках *анализ слов с точки зрения количества в них согласных и гласных фонем* позволяет разделить слова на четыре группы: 1. слова, в которых согласные преобладают над гласными: в АЯ такие слова состав-

* Латвийский государственный университет, Рига.

ляют 58,4 % а в ЛЯ — 56,4 %. В некоторых словах количество согласных достигает в обоих языках 85 %—87 % (в ЛЯ и 90 %), напр., [sɪksəʃs] sixths, [ˈsɪmplɪnz] simpletons; [sprakʃh] sprakšk, [striŋkʃh] strinkšk. 2. слова, в которых число согласных и гласных фонем одинаково составляют в АЯ — 38,7 %, в ЛЯ — 42,2 %; 3. слова, в которых гласные преобладают над согласными, в АЯ составляют 0,5 %, а в ЛЯ — 1,5 %. В некоторых словах в ЛЯ количество гласных достигает 75 %, напр., в слове [ˈne-ie-ei] neieej, в АЯ — 66,6 %, напр., в слове [ˈouɐ] over, [ˈouɪŋ] owing.

4. слова, которые состоят только из гласных фонем в АЯ составляют 2,5 %, а в ЛЯ имеется всего несколько слов, напр., в АЯ: [ɪə] ear, [aʊ] hour; в ЛЯ [èi] ej, [ie-èi] ieej, [äu] au, [ie-ai] ieau.

5. Анализ количественного соотношения согласных и гласных фонем показывает, что как в АЯ и в ЛЯ согласные преобладают над гласными, но соотношение в каждом языке неодинаково: в АЯ согласных больше, чем гласных на 26 %, а в ЛЯ на 16 %.

6. При анализе фонемного состава текстов была определена частотность употребления отдельных групп фонем и отдельных фонем: 1. В обоих языках число звонких согласных превышает число глухих согласных: в АЯ звонкие согласные составляют 66 %, глухие — 34 %; в ЛЯ звонкие согласные составляют 60 %, а глухие — 40 %. Следовательно, процентуальное соотношение почти одинаково. 2. В обоих языках шумные согласные употребляются больше, чем сонанты: в АЯ шумные составляют 65 %, а сонанты — 35 %, а в ЛЯ шумные составляют 77 %, а сонанты — 23 %. Такое соотношение в употреблении шумных и сонантов в каждом языке свидетельствует о большей сонорности АЯ по сравнению с ЛЯ. 3. В ЛЯ смычные согласные (54 %) преобладают над щелевыми (46 %), а в АЯ наоборот, щелевые согласные (52 %) преобладают над смычными (48 %). 4. В ЛЯ мягкие согласные составляют 8,8 %, а в АЯ 0,5 %. 5. В обоих языках свистящие и шипящие составляют 17 %. 6. В обоих языках монофтонги встречаются чаще, чем дифтонги: в АЯ монофтонги составляют 86 %, а дифтонги — 14 %, в ЛЯ монофтонги составляют 84 %, а дифтонги — 16 %. 7. В обоих языках долгие монофтонги встречаются реже, чем краткие монофтонги, но процентное соотношение неодинаково: в АЯ долгие монофтонги составляют 12 %, а краткие — 88 %, а в ЛЯ долгие монофтонги 32 %, а краткие 68 %. 8. При сравнении частотности отдельных фонем оказывается, что в АЯ преобладает монофтонг [ə], употребление которого составляет 11 % по отношению ко всему фонемному составу АЯ, а по отношению к гласным АЯ — 30 %; за [ə] следует гласная краткая фонема (ɪ), употребление которой составляет 9 % и 25 %; за (ɪ) следуют (e), (æ), (ʌ), (v), (i:), (ɔ:), (a:), (v), (z:). В ЛЯ чаще всего встречается гласная фонема (a) — 10 % и 25 %, потом следуют (a:), (i), (u), (e). 9. Из инвентаря дифтонгов АЯ чаще всего встречается (aɪ), за ним (eɪ), (oɪ), (aʊ) а в ЛЯ: (uo), (ie), а за ним следуют (ai), (au). 10. Из инвентаря согласных

фонем в АЯ чаще всего используется (n) — 6,5 % и 10 %, затем следуют (t), (s), (θ), (d). В ЛЯ чаще всего используется согласная фонема (s) — 7 % и 13 %, потом следуют (r), (t), (k), (n).

7. Особенности распределения гласных фонем в слове в АЯ и ЛЯ во многом неодинаково, но имеются и общие особенности: в ЛЯ все монофтонги и дифтонги могут стоять как в закрытых, так и в открытых слогах в слове, но в АЯ краткие монофтонги (e), (æ), (ʌ), (v) могут находиться только в закрытых слогах.

В ЛЯ все гласные фонемы, за исключением (æ) и (æ:) могут стоять в конце слова в открытом слоге, а в АЯ в этом положении не могут стоять не только (æ), но и (e), (ʌ) и (v). Имеются еще и другие различия.

8. Согласные фонемы распределяются в словах в обоих языках во многих отношениях одинаково, но имеются и различные особенности, например, не все из согласных (r), (h), (ʒ), (j), (v) и (w) могут стоять как в начале, так и в конце слова. В АЯ (r), (h), (j) и (w) не могут стоять в конце слова / (r) произносится в конце слова, если следующее слово начинается с гласной/. В ЛЯ согласные (h), (ʒ), (j) и (v), не могут стоять в конце слова, напр., (ʒ) стоит в конце слова только в слове (pɪʒ-pɔʒ) ping-pong, а (r) и (j) произносятся в конце слова только тогда, если следующее слово начинается с гласной, а в остальных случаях они вокализуются в (u) и (i) и вместе с предыдущей краткой гласной образуют истинные дифтонги. Имеются еще и другие особенности.

9. Сочетание фонем в различных положениях в слове имеют много общих особенностей, в обоих языках, но имеется и много различных особенностей:

1. в обоих языках в начале слова могут сочетаться две гласные фонемы напр. в АЯ: [ˈaɪədaɪn] iodine; в ЛЯ: [ˈie-ai-jä:] ieaijä.

Также две гласные могут сочетаться в обоих языках как в конце, так и в середине слова. Напр., в АЯ: [ˈʃou-ɪ] showy; в ЛЯ: [ˈnɔ-èr] noej, [ˈsa-aüg] saug. В ЛЯ в конце слова могут сочетаться не только две, но и три гласные фонемы: [ˈne-ie-ei] ne ie ej.

2. В сочетаниях согласного с гласным в ЛЯ имеются слова, в которых после начального согласного сочетаются три гласные фонемы, напр., [ˈjâ:-ie-ai-ci-na] jaieaicina. Такое явление не наблюдается в АЯ.

3. В обоих языках слова могут начинаться с сочетанием двух и трех согласных (напр., (tr), (pl), (str), (spr), (spl), однако, в конце слова в ЛЯ больше четырех согласных не сочетается (напр., [striŋkʃh] strinkšk, [taŋkʃh] tarkšk, но в АЯ имеются слова, в которых сочетаются даже пять и шесть согласных фонем (напр., [ˈfʌŋkʃnz] functions, [ˈsɪmplɪnz] simpletons).

4. Состав согласных фонем в сочетаниях в обоих языках во многих случаях один и тот же: (sn), (skr), (spl), (spr), но имеются сочетания, в которых состав и количество согласных фонем различны. 5. В обоих языках имеются

сочетания согласных фонем, которые используются как в начале, так и в конце слова, но имеются и такие начальные сочетания, которые в конце слова не используются, а некоторые конечные сочетания не встречаются в начале слова.

Нами исследованы также все сочетания фонем возможные, в обоих языках.

DISCUSSION

Schneider-Kobersky:

1. Has your investigation been carried out on written or oral material?
2. Are your conclusions based on connected speech or isolated words?
3. Would you consider, for instance, in the expression "the apples are red" [tʒ] as a consonant cluster?

Vecozol:

Ad Schneider-Kobersky:

1. My investigation has been carried out on transcribed material—different kinds of texts (fiction: prose and dialogues, and technical texts).
2. My conclusions are based on the analysis of words in connected speech (as the texts were read aloud and transcribed).
3. I consider in the expression "the apples are red" [b₁læptʒəred]—the consonant combination [ptʒ] as a three—consonant cluster.

АКУСТИЧЕСКАЯ ХАРАКТЕРИСТИКА ВОСПРИНИМАЕМЫХ ЗВУКОВЫХ ЕДИНИЦ РУССКОЙ РЕЧИ

Л. А. ВЕРБИЦКАЯ*

Целью настоящей работы было выяснение вопроса о том, замечают ли носители языка те звуковые различия, которые не имеют фонематического значения, т. е. различия между оттенками одной фонемы.

Были взяты все ударные и безударные гласные во всех возможных положениях. Для анализа, а также для опытов по восприятию гласные выделялись из слов на слух с помощью особого прибора „магнитофона с вращающимися головками“ при обязательном осциллографическом контроле. Спектральный анализ производился на спектрометре типа „видимая речь“. Были сняты спектрограммы выделенных гласных и исходного материала в произношении 7 дикторов. Весь материал был прослушан 50 аудиторами.

Анализ полученных от аудиторов данных показал, что различаются 18 эталонов. Эталоны № 1—№ 4 связаны с фонемой *a*, № 5—№ 8 — с фонемой *o*, № 9—№ 12 — с фонемой *u*, № 13—№ 16 — с фонемой *e*, № 17 — с фонемой *i*, № 18 — с фонемой *ы*. Все эталоны можно разделить на 4 группы:

1. эталоны, соотносящиеся с оттенками гласного в слоге типа *tat* для *a*, *o*, *u*, *e*, *ы*; для *i* в слоге типа *t'at*
№ 1а, № 5о, № 9и, № 13е, № 18ы, № 17i
2. эталоны, соотносящиеся с оттенком гласного в слоге типа *ta't* № 2а', № 6о', № 10и', № 14е'
3. эталоны, соотносящиеся с оттенком гласного в слоге типа *t'at* № 3'а, № 7'о, № 11'и, № 15'е
4. эталоны, соотносящиеся с оттенком гласного в слоге типа *t'at'* № 4'а', № 8'о', № 12'и', № 16'е'

Установленный набор эталонов показал, что аудиторами замечаются только различия, связанные с мягкостью или твердостью соседних согласных; различия же, которые определяются соседством с согласными, отличающимися способом образования, участием голоса, назализацией, действующим органом, аудиторами не воспринимаются.

Исследование акустических характеристик эталонов всех групп показа-

* ЛЭФ Ленинградского университета, Ленинград.

ло, что для всех *a*, отождествленных с эталоном № 1*a*, характерно усиление в области около 1000 гц, для всех *o*, идущих в эталон № 5 *o* — усиление полосы 300—800 гц; для всех *u*, соотносимых с эталоном № 9 *u* — полосы 100—500 гц, для всех *e*, отождествляемых с эталоном № 13*e* — характерно усиление 2-х полос: до 500 гц и от 1300—1700 гц; для всех *o*, соотносимых с эталоном № 17 — тоже 2-х полос: до 500 гц и около 2000 гц; для *y*, — тех же полос. Для эталонов 1-й группы характерно ровное положение F_{II} . Для эталонов 2-й группы — восходящее движение F_{II} , от частоты, характеризующей соответствующие гласные, до 2000 гц; для эталонов 3-й группы — нисходящее движение F_{II} от 2000 гц — до частоты, характеризующей соответствующие гласные; для эталонов 4 группы — или нисходяще-восходящее движение (от 2000 гц — до частоты, характеризующей соответствующие гласные и до 2000 гц) или вообще более высокое ее положение.

Исследование показало, что предъявленный аудиторам гласный отождествлялся с тем эталоном, со спектральными характеристиками которого совпадают его спектральные характеристики. В большинстве случаев все предъявленные *a*, например, были восприняты как *a*, отождествлены с эталоном № 1*a*; все *u* — с эталоном № 9*u* и т. д. Однако имеются случаи смешения оттенков одной фонемы. Так, 20 % всех звуков, воспринятых как *a*, в действительности соответствуют — *a'*, 25 % *o* — *o'*, 31 % *u* — *u'*, 17 % *e* — *e'*. Это означает, что *i* — образный переход, появляющийся в звучании гласного, соседнего с мягким согласным, в таких случаях либо вовсе отсутствует, либо по каким-то причинам не воспринимается носителями языка. Такими причинами могут быть или малая длительность перехода, или слишком небольшое различие в звучании по сравнению с характерной частью. Анализ спектрограмм тех гласных, которые были восприняты как соответствующие гласные в положении между твердыми, показал, что *i* — образный переход в их спектре отсутствует.

Кроме того, не исключено отождествление с одним и тем же эталоном оттенков разных фонем, смешение этих разных фонем. Так, 10 % всех звуков, воспринятых как *e*, в действительности соответствуют фонеме *a* в положении между мягкими (эталон № 13*e*), 16 % всех звуков, отождествленных с *i* — фонеме *e* в положении между мягкими (эталон № 17*i*). Рассмотрение спектрограмм *a*, воспринятых как *e*, и *e'*, воспринятых как *i*, свидетельствует о том, что такое восприятие не было случайным: области концентрации энергии таких *a'* и *e*, *e'* и *i* совпадают.

При восприятии безударных гласных носители русского языка руководствуются набором эталонов ударных, однако, используя только 6 из них: № 1*a*, № 9*u*, № 11'*u*, № 13*e*, № 17*i*, № 18 *y*.

Исключительно важной оказывается роль временного фактора: при длительности гласного менее 80 мсек соотношение предъявленного оттенка фонемы с соответствующим эталоном становится затруднительным.

Таким образом, можно утверждать, что носители языка замечают не только различия звуков, противопоставленные фонематически, но и различия, которые фонематического значения не имеют. Набор звуковых типов, хранящихся в памяти человека, не ограничивается числом фонем, а превышает его; это превышение в русском языке происходит за счет тех оттенков гласных, которые возникают в соседстве с мягкими согласными; последовательность, с которой аудиторы отмечают эти оттенки, заставляет предположить, что они занимают принципиальное иное положение, чем остальные оттенки. Эта разница заключается в особой функциональной нагрузке гласных в соседстве с мягкими согласными, мягкость согласного наиболее заметно и последовательно реализуется в соседнем гласном. „Перераспределение“ нагрузки существенного для фонологической системы русского языка, признака приводит к тому, что гласные, выступающие в окружении мягких согласных, образуют функционально выделенную группу оттенков.

Таким образом, выявившиеся 18 эталонов являются „типовыми“ гласными, которые хранятся в памяти носителей русского языка и с которыми закономерно соотношены оттенки фонем, а через них и фонемы.

DISCUSSION

Wiede:

Die Untersuchungsergebnisse haben eine große Bedeutung für die Methodik des Fremdsprachenunterrichts, besonders bezüglich der Aneignung einer guten Aussprache. Dem Vortrag konnte leider nicht entnommen werden, ob die untersuchten Vokale mit allen im Russischen vorhandenen Konsonanten kombiniert worden sind. Es ist anzunehmen, daß die einzelnen Konsonanten unterschiedlich auf die Qualität der Vokale einwirken.

THE RELATION BETWEEN THE OBJECTIVE FREQUENCY OF THE FUNDAMENTAL TONE AND THE SUBJECTIVELY PERCEIVED PITCH AND MELODY OF THE OESOPHAGEAL VOICE

KAREL VRTIČKA*

The substitute esophageal voice is generally considered as the best compensation of alaryngeal aphonia after total laryngectomy. (The basic importance of the upper esophageal mouth as a vicarious glottis was recognized by Seeman as early as 1919; this author was also the first to describe the physiology and pathophysiology of the substitute esophageal phonation.) The sound of the esophageal voice is highly adaptable and serves as a basis for a fluent and intelligible substitute speech with its natural melody, accents and personal vocal timbre.

193 tape recordings of free narrating from 116 superior oesophageal speakers (rehabilitated at the Phoniatic University Clinic in Prague) were chosen for this study; the group consisted of 112 men and 3 women. The results of subjective examination of vocal pitch were correlated with objective sonographic analyses.

The fundamental frequency was determined in 113 esophageal speakers. Individual mean frequencies of 109 male speakers ranged from 21.42 cps to 186.00 cps (average frequency 63.57 cps) and of 4 female voices from 98.88 cps to 202.50 cps.

There is, however, a substantial difference between the objective acoustical structure of the voice and its subjectively perceived qualities. According to Sedláček and Sychra, the fundamental tone is not perceived separately and the acoustical components of speech are not analysed one by one, but the global sound of speech as a whole is evaluated by the listener. The perception of the melody of speech and the recognition of vocal timbre are thus attached together and form an inseparable complex of vocal timbre and pitch.

When comparing the results of previous subjective examination of vocal pitch with the objectively measured and calculated mean fundamental frequencies of individual esophageal voices, we very often noticed a striking discrepancy. Higher frequency and intensity prevalence of the second—and sometimes also of the third and fourth formants—resulted in a markedly bright vocal timbre causing a subjectively perceived high pitch of the voice. Our sonagrams demonstrate this fact very clearly. Because of higher formant frequencies and a marked intensity prevalence of

* From the Phoniatic Research Institute, Medical Faculty, Charles University in Prague (Director: Prof. Dr. M. Seeman, DrSc.).

superior formants, the esophageal voice in Fig. 1 was subjectively perceived as high-pitched, even though its fundamental frequency of 38,11 cps is one octave lower than that of 83,14 cps of the esophageal voice in Fig. 2 which was subjectively perceived as much lower in pitch because of the very low frequency as well as of a marked intensity prevalence of the first formant.

According to Van den Berg and Moolenaar-Bijl as well as Tato and al., the subjectively perceived changes of vocal pitch depend much more upon changes of vocal timbre than on actual changes of fundamental frequency of the esophageal voice.

In 97 esophageal speakers of our series, a comparison of previous subjective examination of pitch changes with the objective measurements of frequency changes—syllable by syllable—was possible. Again, a marked discrepancy was very frequently stated.

Only in 45 speakers, the subjectively perceived pitch variations corresponded to the objectively measured changes of fundamental frequency. Complete agreement was found in 30 cases (as e.g. in Fig. 3), while in the remaining cases, the sense of the variations agreed but the subjectively perceived pitch variations were much greater than the changes of frequency. Correspondence of the subjectively perceived pitch variations and the objectively measured changes of frequency was most frequently found in esophageal voices with relatively high fundamental frequencies (average frequency of this group was 77, 78 cps). On the contrary, the pitch variations did not correspond to the changes of frequency in 52 speakers. In 18 cases, no changes of frequency (except the final decrease on the last syllables of the sentence) were found in the examined utterings. A gradual decrease of frequency—which, however, did not correspond to marked pitch variations—was found in 26 cases and, in the remaining 8 cases, both the pitch and frequency varied but in contrary senses. The subjective impression of increasing or decreasing vocal pitch depended, in the latter 52 cases, almost exclusively upon the changes of the formant pattern of the esophageal voice resulting from intentional brightening or darkening of vocal timbre. Fig. 4 illustrates this fact. Both of the melodic tops of the sentence correspond to syllables displaying the deepest decrease in fundamental frequency; a marked increase in all formant frequencies as well as an upward shift of the sound energy is characteristic of these syllables, however. Average frequency of the 52 voices of the second group was 51,92 cps (the difference is statistically significant).

We may thus conclude that the subjective perception of pitch variations in the course of fluent esophageal speech substantially depends upon two respective factors. In esophageal voices with higher fundamental frequencies, the subjective impression of pitch variations depends mainly upon the changes of the fundamental frequency. In esophageal voices with low and very low fundamental frequencies, the subjective impression of pitch variations corresponds mainly to the variations of vocal timbre.

DISCUSSION

Lehiste:

I was interested in the difference in the average fundamental frequencies of the esophageal voices of men and women. Is there a physiological reason, or is there a psychological reason, the women making a greater effort to reach a socially acceptable pitch level?

Sovijärvi:

Could you briefly tell us which are the cues of the timbre changing the subjective melody patterns of oesophageal voice?

Handzel:

Manche unserer Kehlkopflosen behielten nach der Rehabilitierung die regionale Sprache, in der die Veränderung der Intonation während des Sprechens, z. B. des Wortes „Lwow“ (als „Lwowa“ oder „Lwowie“) subjektiv perzeptiert wird. Auf Grund unserer Forschungen tritt hier nicht nur eine Veränderung der Höhe von einzelnen Formanten ein, sondern auch der Höhe des Grundtons, also der Melodie.

Vrtička:

ad Lehiste: I do not think there is a substantial difference between male and female substitute glottis responsible for higher fundamental frequencies of female esophageal voices. However, a high-pitched esophageal voice—which is socially more acceptable—was trained intentionally in all our female laryngectomies. On repeated tape recordings and sonographic analyses in one of them, we were able to demonstrate that the average fundamental frequency rose, in the course of vocal rehabilitation, from initial 49,00 cps to 98.88 cps.

ad Sovijärvi: According to our findings, the subjective impression of increasing or decreasing vocal pitch results from intentional brightening or darkening of the vocal timbre of the esophageal voice i.e. from shifting the resonance bands and changing the respective intensities of formants (shifting the sound energy maxima) by means of exaggerated articulation movements.

Vrtička: The Relation between the Objective Frequency of the Fundamental Tone and the Subjectively Perceived Pitch and Melody of the Nesophageal Voice

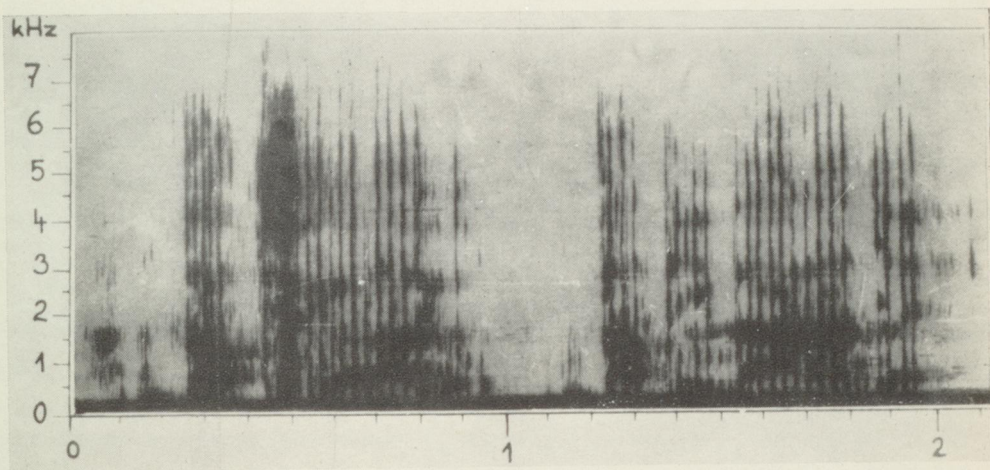


Fig. 1.

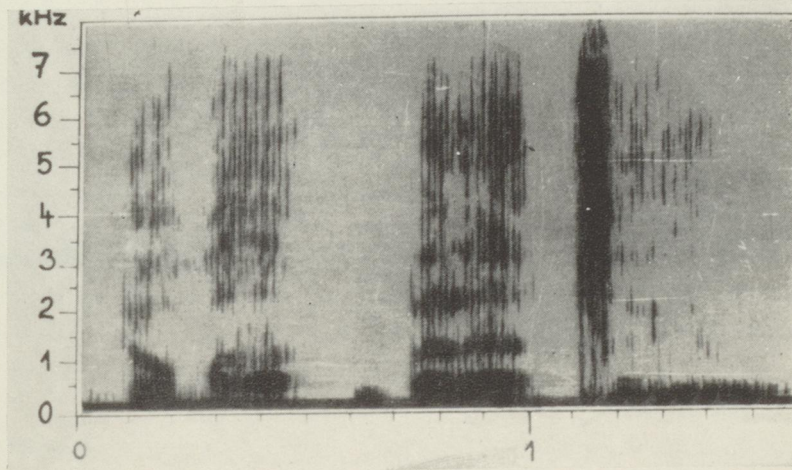


Fig. 2.

Vrtička: The Relation between the Objektiv Frequency of the Fundamental Tone and the Subjectively Perceived Pitch and Melody of the Oesophageal Voice

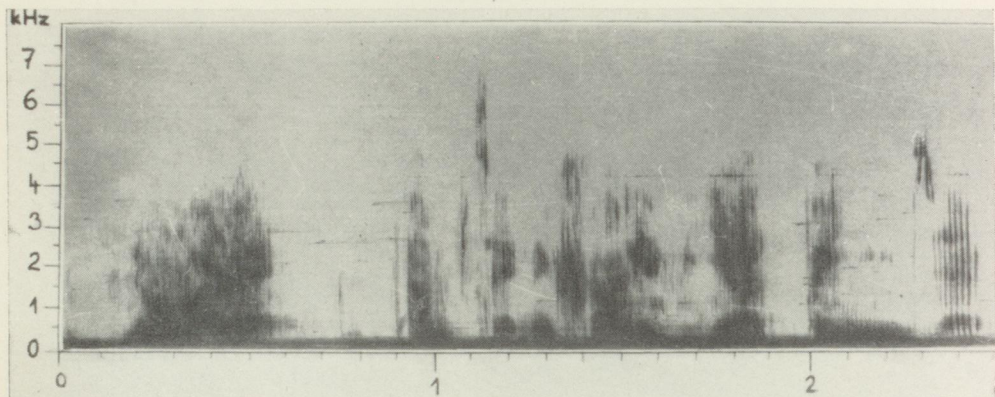


Fig. 3.

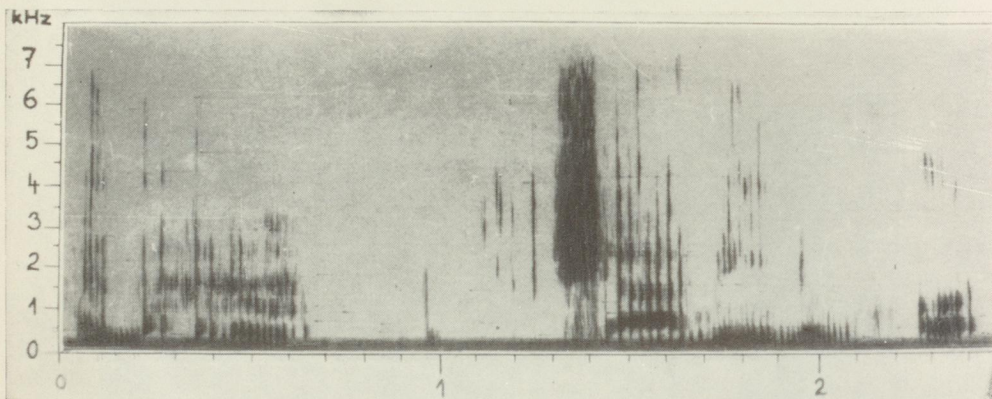


Fig. 4.

IDENTIFICATION DE VOYELLES EN FONCTION DE LEUR DURÉE

M. WAJSKOP*

Cet exposé décrit les résultats préliminaires obtenus au cours d'une série d'expériences destinées à mesurer les seuils d'identification de voyelles isolées en fonction de leur durée.

Ce point de départ peut paraître contestable à des linguistes et à des phonéticiens. Les segments sonores isolés ont fait, durant le premier quart de ce siècle, l'objet de recherches intensives mais ont été dédaignées, soit que l'appareillage utilisé fût médiocre, que les résultats fussent mal présentés ou que le montage expérimental négligeât les données élémentaires de la psychologie. De ce fait, la plupart des travaux récents partent d'unités considérées comme naturelles. On peut cependant se demander s'ils n'ont pas envisagé le problème de la perception de la parole à un niveau trop élevé. En effet, comme nous savons peu de choses sur la manière dont le signal acoustique est déchiffré, l'analyse qui s'impose au préalable doit vérifier la présence de tous les éléments redondants contenus dans le signal (forme spectrale, qualité de la voix, indices temporels). C'est dire qu'il faut aborder auparavant ce stade d'une perception que l'on peut qualifier de primaire où les segments sonores seraient appréhendés, ainsi que l'a noté Öhman, suivant un processus de consécution simple.

Protocole expérimental: 3 voyelles isolées [i], [a], [u], prononcées par un homme et une femme ont été enregistrées. De ces énoncés, normalisés en intensité et contrôlés auditivement et spectrographiquement, on a réalisé, à l'aide d'une porte électronique, des segments de 4 à 400 ms. Deux bandes-tests de 60 stimuli (3 voyelles x 20 durées) disposés en ordre aléatoire ont été administrées individuellement à 30 sujets. Ceux-ci avaient le choix entre 4 réponses (i — a — u — zéro).

Expérience I

I.A: 10 sujets furent soumis au test de la bande n° 1 (voix d'homme). Des résultats, il ressort un % d'erreurs de 4,07. Cependant, on constate que la totalité des erreurs se concentre au-dessous de 20 ms et que, même à 4 ms, le score de reconnaissance atteint 0.77.

* Université de Bruxelles.

I.B.: A 10 autres sujets fut administré le test de la bande n° 2 (voix de femme). Le % d'erreurs fléchit à 2,03. Ici également, les erreurs se concentrent au-dessous de 20 ms et le score à 4 ms atteint 0.80.

I.C.: En fonction des résultats obtenus antérieurement, les durées supérieures à 40 ms ont été éliminées et les deux voix ont été mêlées sur une nouvelle bande magnétique comprenant 72 stimuli (2 x 3 voyelles X 12 durées). Les résultats indiquent une montée du % d'erreurs qui atteint 9,86. Le score de reconnaissance à 4 ms tombe à 0.67. Si l'on sépare les 2 catégories de stimuli, Vm et Vf, et que l'on considère que le groupe de 10 sujets a fonctionné différemment selon ces 2 classes, on obtient pour Vm un % d'erreurs de 13,3 (360 observations) et pour Vf, un % d'erreurs de 6,4.

COMMENTAIRES

- (I) Validité des résultats: Les réponses des sujets sont toujours supérieures à la limite du hasard; la zone d'incertitude est réduite et les réponses atteignent rapidement un score de reconnaissance élevé.
- (II) Score de reconnaissance en fonction de F_0 : La meilleure identification des voyelles prononcées par la voix de femme (pour l'ensemble des expériences, Vm: 9,72% d'erreurs — Vf: 4,72%) semble bien être due au fait que l'on offre à l'auditeur, pour un même laps de temps, un échantillon d'informations plus grand.
- (III) Score de reconnaissance en fonction de la durée: Nous avons étudié la distribution des réponses correctes en fonction des erreurs commises pour chaque condition de durée et pour chaque voyelle.

L'indice d'identification est remarquablement élevé même aux très courtes durées. Exemples:

	a	u	i
4 ms	0.73	0.83	0.63
6 ms	0.75	0.90	0.90
8 ms	0.90	0.88	0.90

Ce résultat doit cependant être considéré avec prudence étant donné l'hétérogénéité des observations ainsi rassemblées et la limitation du vocabulaire soumis aux sujets.

Expérience II.

Nous nous sommes demandés quelles seraient les réactions d'auditeurs si, placés en face de ces mêmes stimuli, il leur était permis de ventiler leurs réponses sur un choix plus large, en l'occurrence les 7 voyelles du système français minimal (1): [i], [u], [y], [A], [E], [OE] et [O]. Cette expérience ne concerne jusqu'à présent que

10 sujets. Les 3 bandes-tests ont été normalisées aux 12 durées initiales. La matrice globale indique la direction des confusions.

	A	i	u	OE	E	O	y	
A	452	1	3	7	13	4	0	(47%) 480
i	0	422	10	1	30	4	13	(52%) 480
u	7	6	424	2	3	34	4	(61%) 480
	459	429	437	10	46	42	17	440

Fréquence des réponses correctes: 1298 — score d'identification 0.901

Les tendances sont claires: [A] et [i] tendent à être confondus avec [E] tandis que [u] glisse vers [O]. Les voyelles fermées sont donc entendues plus ouvertes, soit moins aigüe, soit moins grave. La voyelle ouverte et centrale est entendue plus fermée et plus antérieure.

Si l'on transcrit graphiquement, (fig. 1) sous forme de résultantes de vecteurs, les tendances ainsi exprimées et ce, à l'intérieur du triangle acoustique décrit par Delattre (2), on observe que les directions des 3 résultantes se croisent dans un espace exigü, proche de la région du schwa. Ceci confirme sur le plan perceptif mais de manière indirecte et non quantitative, les conclusions énoncées par Lindblom (3) sur le plan articulatoire et acoustique.

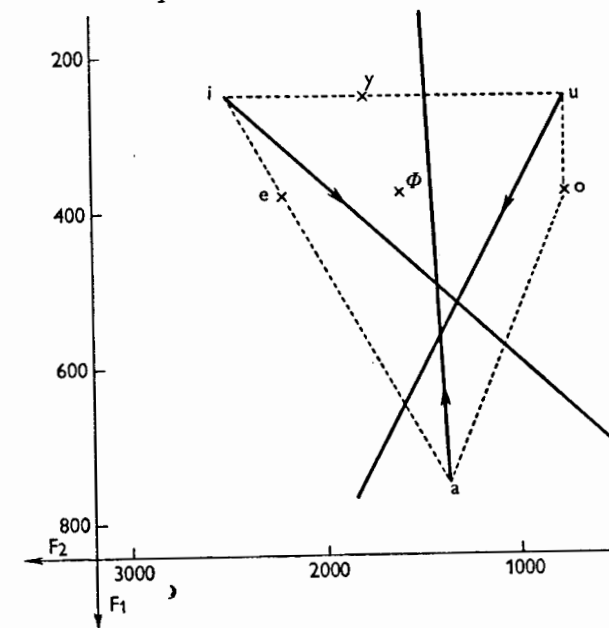


Fig. 1.

CONCLUSIONS

1. Sur le plan de l'identification, nos résultats confirment ceux déjà obtenus par Gray (4), Peterson (5) et Cramer (6): reconnaissance aisée de vocoides dont la durée n'atteint pas une période complète.
2. (i) La réduction temporelle, à l'intérieur d'un cadre expérimental symétrique, tend à confondre les voyelles les plus fermées (i et u), à perturber la reconnaissance de [a] ou à la confondre indifféremment avec les voyelles fermées d'avant ou d'arrière.
(ii) L'augmentation de la hauteur tonale améliore les performances des sujets.
3. (i) Par contre, à l'intérieur d'un cadre asymétrique, les voyelles fermées sont entendues plus ouvertes dans leur plan respectif tandis que la voyelle centrale tend à se diriger vers le haut et l'avant.
(ii) Les scores d'identification pour les courtes durées sont inférieurs à ceux obtenus dans la première série d'expériences.

REFERENCES

1. B. Malmberg, La structure phonétique de quelques langues romanes. *Orbis*, XI, 1, 131—178 (1962).
2. P. Delattre, Comparing the vocalic features of English, German, Spanish and French. *Iral*, II, 2, 71—97 (1964).
3. B. Linblom, On Vowel Reduction. *Report n° 29, STL, RIT*, Stockholm (1963).
4. G. W. Gray, Phonemic Microtomy: the minimum duration of perceptible speech sounds. *Speech Monographs*, IX, 75—90, (1942).
5. G. Peterson, The Evaluation of Speech Signals. *JSHR*, XIX, 2, 168 (1954).
6. B. Cramer, Über das Erkennen von Sprachlauten. *Aufnahme und Verarbeitung von Nachrichten durch Organismen*, pp. 63—70 (1961).
7. S. Öhman, On the Perception of Swedish Consonants in intervocalic position. *Report n. 25, STL, RIT*, Stockholm (1962).

DISCUSSION

Rossi:

Le score de reconnaissance très élevé dans votre première expérience est dû à la différence de hauteur des voyelles présentées aux sujets; si cette différence de hauteur n'est pas spécifique de la voyelle, elle doit évidemment être supprimée; mais ne doit-on pas tenir compte de la hauteur spécifique (voir les travaux de Fairbanks et Peterson) qui fait partie intégrante des propriétés des voyelles et présenter par conséquent les voyelles à des hauteurs différentes? Ne risque-t-on pas, dans le cas contraire, de fausser les résultats de l'analyse? Ma remarque vaut également pour l'intensité spécifique.

Wajskop:

Les stimuli ont été normalisés afin de répondre à une demande du département de psychologie expérimentale. Dans les expériences ultérieures, il a été tenu compte des différences spécifiques de hauteur (Peterson-Lehiste) et d'intensité (Fairbanks) afin d'éliminer le caractère "synthétique" de nos stimuli. Dans la Ière expérience, les confusions ne se dirigent pas nécessairement vers les voyelles ouvertes quand il s'agit de [i] et [u]. On note des confusions dans le sens horizontal de [i] vers [u] et l'inverse.

PERCEIVING COMPLEX SPEECH AND NON-SPEECH SOUNDS

J. C. WEBSTER*—A. CARPENTER**—M. WOODHEAD**

The purpose of this paper is to see if naive listeners can learn supposedly meaningless sounds quicker if some are indeed sustained vowels, some musical instrument tones and some truly meaningless. Nine sounds were chosen; three were identifiably different but meaningless buzzes; three were two-second portions taken from the middle of the sustained vowels a, i and u; and three were mid-portions of sustained tones on the clarinet, the cello and the tromba- stop on an experimental electronic organ. The subjects were not told that there were three different classes of sounds although they were told "... some may sound like voice sounds, some... musical... (or) perhaps they will (all) sound like nothing you've heard before".

STIMULUS MATERIALS

A 54-item sound identification (*SI*) test was recorded on magnetic tape. Each item consisted of four tonal complexes. The duration of each tone, which could be any one of nine sounds, was two seconds. Each interval between tones was also two seconds. A five second interval was left between items to give the listener time to write a response. The vowels were intoned for 15 seconds at a fundamental frequency of 110 Hz by a native British speaker; the two second samples were chosen from the middle of the 15-second utterance. The musical tones were originally recorded at 220 Hz for 15 seconds but played back at half speed for the two-second mid-portion samples used in the recorded test. The meaningless sounds were buzzes with a fundamental frequency of 110 Hz. They differed in that harmonics 6, 12, 18, and 24 were greatly emphasized in one sound, harmonics 1, 8, and 9 in another and harmonics 1, 8, 16, and 24 in the third. Every effort was made to have all nine sounds at the same frequency and intensity on the recorded test.

Two variations were made of the original *SI* test, a "low"-*SIL*-and a "high"-*SIH*-frequency version. For both of them the voice announcements and the five second

* Exchange Scientist from US Navy Electronics Lab., San Diego, California, USA.

** Applied Psychology Research Unit, 15 Chaucer Road, Cambridge, England.

interval between items remained the same. For the "high frequency or SIH" variant the 4-tone test item was played back and re-recorded at double speed, so that all frequencies in all sounds were one octave higher but the duration of tones and intervals between was maintained.

Three tests were made up of nine different complex steady-state sounds. On one test, SIL, all nine sounds had fundamental frequencies of 55 Hz; on test SI all had fundamental frequencies of 110 Hz; and on test SIH all had fundamental frequencies of 220 Hz. Three of the nine sounds were completely meaningless and remained so on all three tests. Three sounds were segments of the intoned vowels a, i and u when the fundamental frequency was 110 Hz. At 55 or 220 Hz the fundamental and all other parts of the complex tones were shifted in frequency one octave and the resultant sounds were perhaps vowel-like but were not a, i and u. Three sounds were the tones of musical instruments when the fundamental frequency was 220 Hz. At 110 and 55 Hz all components were shifted and the resultant tones were music-like at 110 Hz but only musically-derived at 55 Hz.

TEST PROCEDURE

Two groups—groups A and B—of six Royal Navy ratings with normal hearing served as subjects for two 45-minute periods a day for two weeks. During the first week group A listened to test SI where the fundamental frequency—and all formant frequencies—of the vowels were correctly located, while group B listened to test SIL—all nine sounds meaningless. During the second week the groups listened to the test they had not heard the first week.

Subjects were asked to identify the sounds by an arbitrarily assigned number. The only clue given to them was that certain sounds might sound voice-like or like musical instruments, or like nothing they had heard before.

CONCLUSIONS

There were fewer errors in assigning the correct "identity number" to vowel sounds when the frequencies were 110 Hz or "correct". When all sounds were shifted downward one octave errors among the three classes of sounds were equally divided. A further test needs to be run to find out what happens when the musical instrument sounds are at the "correct frequencies".

DISCUSSION

Black:

I wonder whether Dr. Webster is tempted to place a relationship between his results (position to human-like sounds) and a motor theory of speech perception.

Hart:

I am not surprised to hear that the sustained vowel sounds were not identified significantly more easily than the musical sounds, as the vowel sounds, sustained for two seconds are deprived of so many attributes they would have in normal speech, that they can no longer be considered anything other than mere sounds.

Webster:

Ad Black: We showed that steady state vowel sounds with no pitch inflection and not identified as speech sounds were perceived better, than either meaningless or music-like sounds. A present experiment will see if musical sounds are easier than meaningless and speech-like sounds. If musical sounds and speech sounds are equally easy, than I'd say overlearning may explain it. If the speech sounds are easier, then the motor theory has some support.

Author's Note:

Since this paper was presented, the „Music“ (SIH) group has been run. The listeners did not identify musical tones better than other tones. This implies there is something unique in the way listeners perceive speech-like sounds but not in the way they perceive music-like sounds.

DER STIMMLICHE AUSDRUCK IN SPRECHSTIMME UND MUSIK

HERTA WEIHS*

Wir sind gewöhnt, uns vom guten Sprecher einen lebendigen und ausdrucksrichtigen Vortrag, also echten Gefühlsausdruck zu wünschen und schon R. Wagner fordert vom Sänger „Wahrheit im stimmlichen Ausdruck, höchste Reinheit der Sprache, eine der jeweiligen Stimmung entsprechende Klangbehandlung und eine melodische Sprachplastik“; auch W. A. Mozarts Gesangsstil schöpft seine eindringlichste Kraft aus der Sprache, während G. Verdi verlangte „... dann einfach singen, wie das Gefühl es eingibt.“

Um diese Zusammenhänge zu verstehen, sei versucht, die genetischen Merkmale des stimmlichen Ausdrucks in der gesprochenen Sprache und in der Musik aus einer biologisch-entwicklungsgeschichtlichen Sicht zu beurteilen. Die wichtigsten von allen stimmlichen Mitteln, durch die wir unsere innerseelischen Zustände und Vorgänge den Mitmenschen kundgeben, sind nach F. Trojan vor allem die beiden gegensätzlichen Arten von Stimmgebung, welche dem Wechsel von Leistungs- und Ruhehaltung entsprechen: „Kraftstimme“ und „Schonstimme“. Ferner sind es die Bewegungsgrundlagen des stimmlichen Ausdrucks, welche aus der ursprünglichen Einheit von Stimme und gesamtkörperlichen Ausdrucksbewegungen (im Sinne von A. Flach) hervorgegangen sind. Besonderen Ausdruckswert haben auch die Gegensatzpaare von Kopf- und Brustregister sowie von Rachenenge und Rachenweite.

Die „Kraftstimme“ ist der Ausdruck der Leistungshaltung, der ergotropen Funktionsrichtung des vegetativen Nervensystems (im Sinne von W. R. Hess) und der zugehörigen Affekte sthenischen Charakters. Ihre phonetischen Kennzeichen sind verhärtete Stimmensätze und eine stoßweise vor sich gehende aktive Ausatmung. Die Stimm- und Artikulationsmuskulatur ist stark gespannt, die stimmlosen Reibe- und Verschußlaute treten hervor. In dem Plattenbeispiel spricht Käthe Gold Gretchens Gebet von Goethe. Lebensecht drückt sich in der Stimme das martervolle Händeringen, die verzweifelte Bittgebärde Gretchens im seelischen Schmerz aus. Im Sonagramm (Abb. 1) ist der appellierende Aufschrei „Hilf“ wiedergegeben.

Die „Schonstimme“ ist der Ausdruck der Ruhehaltung, der Restitutionszustände

* Aus der Sprachabteilung der H. N. O.-Klinik der Universität Wien. Vorstand: Univ. Prof. Dr. Otto Novotny.

der trophotropen Funktionsrichtung (W. R. Hess). Hier herrschen die weichen Stimmeinsätze mit folgendem Schwellklang vor, die Stimmbandpressung ist mäßig, die Ausatmung fließend und gezügelt. In den Plattenbeispielen: „Nun die Schatten dunkeln“ von E. Geibel (gesprochen von Erik Schumann) und „Bin ich nüchtern, bin ich trunken“ von G. v. Göckingh (gesprochen von Johanna Matz) ist zu hören, wie durch Schonstimmgebung die Nachstimmung, zärtliche Bitte und die sehnsuchtsvolle Bewegung des Körpers zum Partner hin ausgedrückt werden. Das Sonagramm (Abb. 2) zeigt das zärtliche: „Bleibe“ (Johanna Matz).

Was nun den musikalischen Ausdruck betrifft, so bestehen in Analogie zur gesprochenen „Kraft-“ und „Schonstimme“ musikalische Formgegensätze. Im Anschluß an R. Lach, der bekanntlich die gegensätzlichen Mittel des Gefühlsausdrucks in der Musik genetisch auf den Phonnationsakt zurückführt, konnte ich nämlich in experimentellen Untersuchungen feststellen, daß die strukturellen Elemente in der Musik bei ergotropen Gemütsvorgängen auf dem Phonnationsschema der esclamazioni beruhen, bei trophotropen hingegen auf der Phonnationsgestalt der messa di voce basieren.

Die ergotropen Gemütsvorgänge wie z. B. die aggressiven — Trotz, Zorn, Drohung, Fluch — und die defensiven wie Angst, Schmerz sind musikalisch, charakterisiert durch: eine überwiegend auftaktige, zweiteilige rhythmische Akzentuierung; die Melodie zeigt bei sprungweiser Fortbewegung und vornehmlich aufsteigender Tendenz einen starken Auftrieb und abrupten Abfall der Tonlinie, die innerhalb einer weiten Höhenspanne verläuft; bei weiter Harmonie eignet der Melodie eine starke Bewegtheit, das dissonante Element herrscht vor. Hiezu einige Plattenbeispiele: Gesungen von D. Fischer-Dieskau aus Schuberts „Erlkönig“ und „Prometheus“.

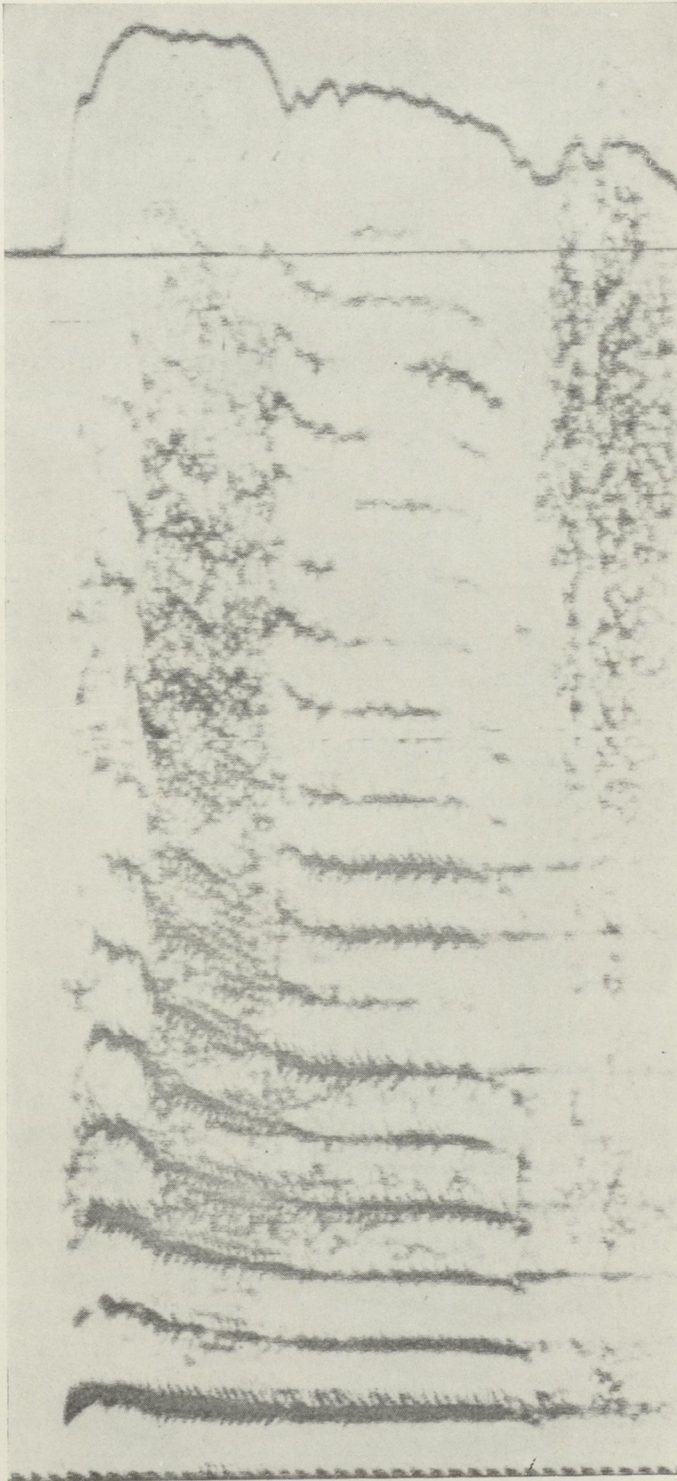
Zunächst das erste Angsterlebnis des Kindes (Takt 41—50): durch die Aneinanderreihung der esclamazioni entsteht das sprunghafte Ansteigen der Melodie, das auch für die sich in der Angst manifestierenden Form der gesprochenen Kraftstimme charakteristisch ist — eine Zackenlinie, die im Umfang einer Sext verläuft. Die Betonung kommt auf die oberen Zacken zu liegen, wodurch ein zweiteiliger motivischer Rhythmus entsteht. Die Drohung des Erlkönigs (Takt 116—123) zeigt wieder echte esclamazioni, die im Zugriff gipfeln. In der folgenden Angstreaktion des Kindes und deren Steigerung zum Schmerzensschrei (Takt 123—131) verengen sich die Intervalle der esclamazioni zu Sekundsritten und die Ausrufe werden durch kurze Pausen unterbrochen, was beides den Ausdruck erhöht. Den Schmerzensschrei wie vorher den Zugriff des Erlkönigs versinnbildlicht Schubert durch einen Quintsprung abwärts. Der stimmliche Ausdruck ist durch Rachenenge und Verstärkung des Kopfklanges in den Angstreaktionen charakterisiert, wogegen in der Drohung und dem Zugriff der Brustklang verstärkt wird. Ähnliches gilt auch für die trotzige Empörung des „Prometheus“: „Hat mich nicht zum Manne geschmiedet...“ mit der Bewegungsgrundlage des Sich-Aufbäumens gegen die überlegene Macht. Schubert charakterisiert sie durch esclamazioni in chromatischer Steigerung (ces, c, cis) auf enharmonischer Basis. Nun zwei Plattenbeispiele des Fluchs aus Verdis „Aida“. Amonasro: „Non sei mia figlia...“

(gesungen von G. Neri); Amneris: „Empia razza... anatema su voi“ (gesungen von M. Pirazzini). Der Bewegungsgrundlage des Zorns und des Fluchs, die auf die Vernichtung des Gegners abzielt, entspricht der Bewegungsablauf des stimmlichen Ausdrucks. Ein Beispiel sardonischer Stimmung in der Musik endlich ist die Serenade des Mephisto aus Gounods „Faust“ (gesungen von Kim Borg). Das mit starker Rachenenge diabolisch klingende Hohngelächter in Oktavsprüngen endet in einer furchtbaren Cacchination.

Nun zu dem musikalischen Ausdruck der trophotropen Gemütsvorgänge, welche dem zum inneren Aufbau gehörigen Ruhegenuß, allen Einswerdungsvorgängen, der Lockung u.ä. nahestehen. Auf der Phonnationsgestalt der messa di voce beruht — im Sinne von Lach — die auf- und absteigende Inflektion der schreitenden Melodie, die sanft fließende harmonische Bewegung sowie eine schwebende Rhythmik. Hiezu Plattenbeispiele aus Schuberts „Erlkönig“, die beiden Lockreden des Erlkönigs. Die erste Lockrede wird ausdrucksrichtig durch Rachenweite und Kopfstimmklang wiedergegeben (D. Fischer-Dieskau) und ruft die Illusion verführerischer Lockung hervor. Sie ist auch durch Lautmalerei von hohem künstlerischen Reiz: im Anfang herrschen I-Laute vor, dadurch glauben wir — im Sinne von Trojan — eine hohe, einschmeichelnde Stimme eines Wesens zu hören, das sich selbst als harmlos und ungefährlich hinstellen möchte. Die Melodie wird hier (Takt 57—71) in dreiteiligem motivischen Rhythmus in sanft an- und abschwellender, pausenfreier Wellenlinie geführt. Die zweite Lockrede (Takt 86—96) bringt einen wiegenden und schwebenden Tanzrhythmus, der durch die sanft fließende Bewegung der zerlegten Akkorde in der Begleitung unterstützt wird.

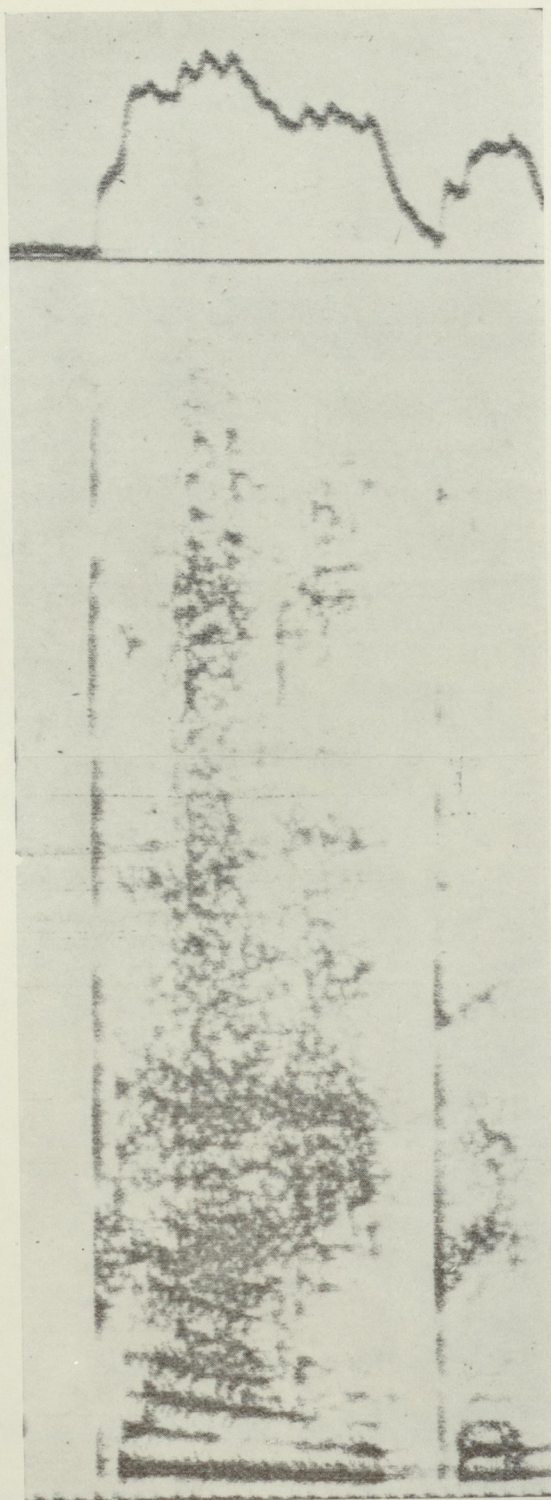
LITERATURANGABE

- R. Lach: Studien zur Geschichte der ornamentalen Melopöie. Kahnt, Leipzig (1913); Die physiologischen Urtypen der musikalischen Formen. *Wien. med. Wschr.* 22 (1927).
F. Trojan: Der Ausdruck der Sprechstimme. 2. Aufl. Maudrich, Wien (1952).
R. Wagner: Der dramatische Gesang. Königsberg 1837.
H. Weihs: Die Beeinflussung der vegetativen Tonuslage durch Sprechstimme und Musik. *Fol. phoniatr.* 619 (1954). Die physiologischen Grundlagen des stimmlichen Ausdrucks. *Österr. Z. f. Musikerz.*, 11. Jg. 1957/58.



Hilf

Abb. 1. Sonagramm: Beispiel für Kraftstimmgebung: Appellierender Hilfeschrei (aus Gretchens Gebet)



Bleibe

Abb. 2. Sonogramm: Beispiel für Schonstimmgebung: Zärtliche Bitte: Bleibe (aus: Bin ich nüchtern...)

ZUR AUDITIVEN STEUERUNG DER SPRECHINTONATION

J. WENDLER*

Der Intonationsverlauf beim Sprechen, hier als Sprechmelodie aufgefaßt, wird meist nach zwei Hauptrichtungen untersucht: einmal als Vermittler logischer Information, also als Satzphonetisches Mittel, zum anderen als Träger emotionalen Ausdrucks. Diese beiden Aspekte sind in unseren Untersuchungen zwei anderen Fragen untergeordnet: Wird die Sprechintonation durch begleitende Musik beeinflußt? Wenn ja, wie wirkt sich diese auditive Steuerung aus?

Als Testmaterial diente uns ein Ausschnitt aus einem Melodram mit streng rhythmisierter, tonhöhenmäßig aber nicht fixierter Sprechstimme. Die Sprechstimme steht in fester rhythmischer Beziehung zur Musik, der Verlauf der Wort- und Satzmelodie ergibt sich aber aus der freien künstlerischen Gestaltung des Sprechers (Fritz Reuter, *Der Hase und der Igel*, nach dem Märchen der Brüder Grimm).

Abb. 1. Der untersuchte Ausschnitt im Melodram.

Unsere ersten Untersuchungen führten wir an mehreren Ausschnitten durch, die unterschiedliche Verknüpfungen von Sprechstimme und Musik aufwiesen. Die gleichen Textstellen wurden bei künstlerisch gestalteter einfacher Lesung und bei melodramatischer Gestaltung (mit Klavierbegleitung) über Kehlkopfmikrofon auf Tonband aufgenommen.

Zunächst arbeiteten wir mit einer Schauspielerin, die das Stück bereits mehrere Male öffentlich aufgeführt hatte. Die entsprechenden Textstellen wurden je 10mal

* Städt. Krankenhaus Prenzlauer Berg, Berlin, HNO-Abteilung — Phoniatrie.

ohne Musik und 10mal mit Musik aufgezeichnet, und zwar an 20 verschiedenen Tagen, um unmittelbare Erinnerungen zu vermeiden. Die Tonhöhenbestimmung der Sprechstimme erfolgte nach zwei von einander unabhängigen Verfahren: einmal auditiv, durch silbenweises Abhören von endlosen Schleifen mit Vergleich am Klavier, zum anderen durch Tonhöhenschreibung, für die wir Herrn Prescher vom Inst. f. Sprechkunde und phonet. Sammlung der Univ. Halle danken. Die nach beiden Verfahren ermittelten Tonhöhen stimmten — innerhalb der methodischen Möglichkeiten — immer überein.

Die Auswertung ergab eine deutliche Beeinflussung der Sprechstimme durch die Musik. Am auffälligsten zeigte sich das an einer Stelle, an der Sprechstimme und musikalische Melodie mit der gleichen rhythmischen Gliederung komponiert waren.

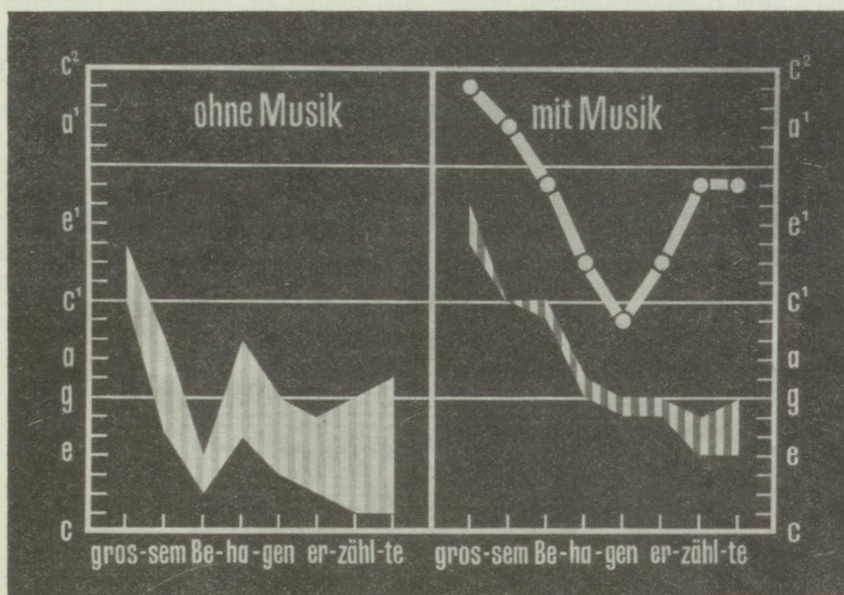


Abb. 2. Mittlere Abweichungen der getroffenen Tonhöhen im 1. Versuch (1 Vpn., je 10 Aufzeichnungen ohne und mit Musik).

Die graphische Darstellung gibt die Streuung der Mittelwerte für die Tonhöhen der einzelnen Silben nach der von Sedláček und Sychra angewandten Methode wieder. Ohne Musik ist die Variationsbreite deutlich größer. Die Anhebung der Stammsilbe des Wortes „Behagen“ kennzeichnet das melodische Profil der Phrase, das entspricht der üblichen hochdeutschen Intonation. Mit Musik streuen die Werte viel weniger, der melodische Verlauf geht in auffälliger Weise parallel zum absteigenden Teil der musikalischen Linie, die in h-Moll von h^1 nach h verläuft. Der vom Komponisten gestaltete Ausdruck des Behagens wird von der Sprecherin übernommen, die Sprechintonation paßt sich ganz der Musik an, auch bei dem Wort „Behagen“. Der progrediente Charakter des Syntagmas, in der musikalischen Linie gut erkennbar, wird nicht realisiert.

Es interessierte uns nun, ob andere Sprecher in ähnlicher Weise reagierten. Wir baten deshalb 30 Berliner Schauspieler, 15 weibliche und 15 männliche Vpn., den Anfang des gelesenen Märchens und des Melodrams für uns zu arbeiten. Wir zeichneten dann wieder dieselbe Stelle auf. Der Zweck der Untersuchung war den Probanden nicht bekannt. Diesmal verfahren wir so, daß jeder Sprecher nur je einmal aufgenommen wurde, einmal ohne und einmal mit Musik, für die Musik benutzten wir Kopfhörer, so daß ein Luftschallmikrophon eingesetzt werden konnte. Die Tonhöhenbestimmung erfolgte wieder auditiv und durch Melographie.

ERGEBNISSE

Zunächst die weibl. Vpn. Bei der reinen künstlerischen Erzählung erkennt man wieder ein Intonationsmuster, das dem Gebrauch in der deutschen Hochlautung entspricht. Im Melodram ist eine Annäherung an die auditiv aufgenommene musikalische Linie

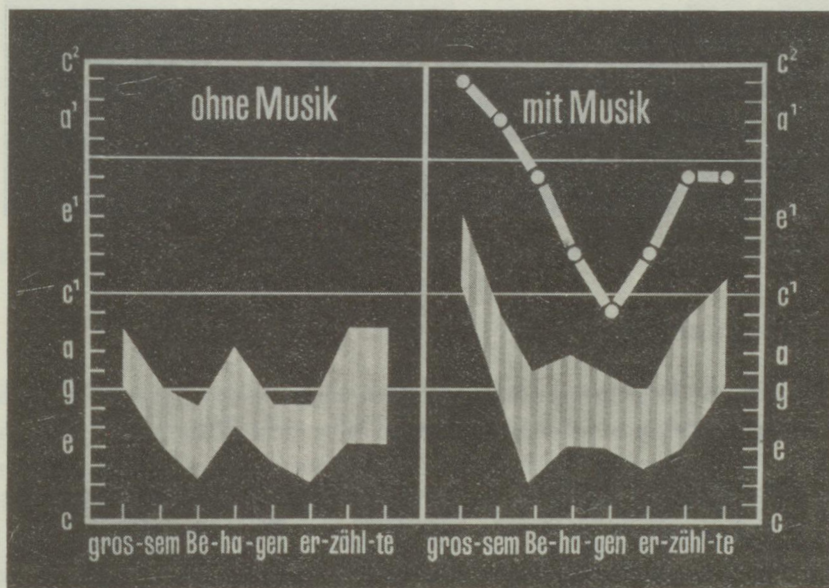


Abb. 3. Mittlere Abweichungen der getroffenen Tonhöhen bei 15 weiblichen Vpn.

unverkennbar, der Grundzug der Intonation ohne Musik ist aber noch angedeutet, das Wort „Behagen“ bricht aus der Parallelität zur Musik aus. Das Ende der Phrase folgt dann wieder der musikalischen Melodie und hebt den weiterweisenden Charakter noch hervor.

Bei den männl. Vpn. zeigt das Profil der Erzählung ebenfalls den üblichen Verlauf, die weitere Streuung der mittleren Abweichungen erklärt sich durch größere Unterschiede in den Stimmlagen der Probanden. Um so auffälliger ist dann aber die Ein-

engung beim Sprechen mit Musik. Dabei zeigt das Feld der getroffenen Tonhöhen wieder eine deutliche Anpassung an die musikalische Linie, das ursprüngliche Intonationsmuster tritt nicht mehr in Erscheinung, die Sprechmelodie folgt der Musik vom Anfang bis zum Ende.

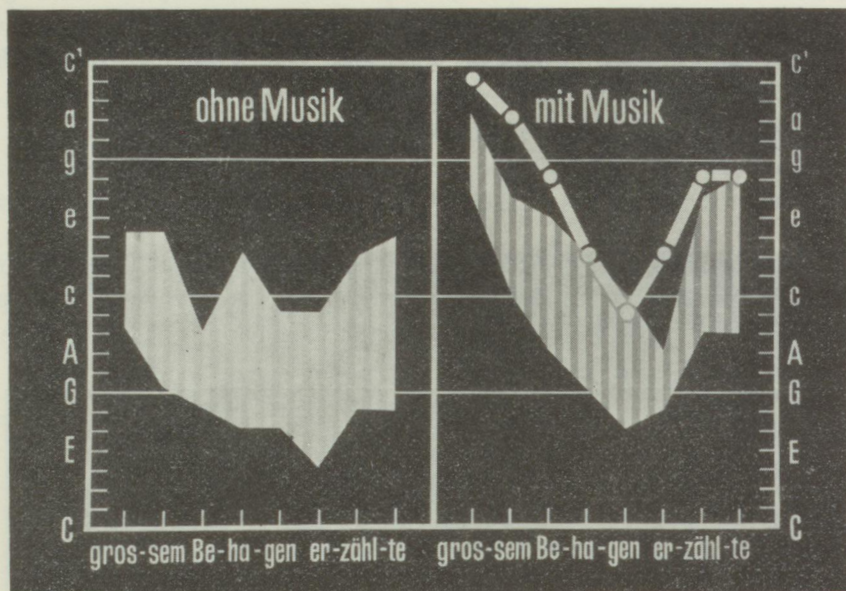


Abb. 4. Mittlere Abweichungen der getroffenen Tonhöhen bei 15 männlichen Vpn.

SCHLUSSFOLGERUNGEN

Die musikalische Begleitung setzt in dem von uns gewählten Ausschnitt als akustisches Signal gleichzeitig zu jeder gesprochenen Silbe einen Hörreiz. Die Tonhöhen-einstellung der Stimme wird dadurch deutlich und gleichsinnig bei allen Vpn. beeinflusst. Die audiophonatorische Kontrolle wirkt sich aber nicht auf direkt-reflektorischem Wege vom Ohr auf den Kehlkopf aus, etwa als unmittelbare frequenzmäßige Steuerung. Die individuellen Unterschiede im Verlauf der sprechmelodischen Kontur zeigen, daß die gesamte psychophysische Persönlichkeit des Sprechers bei der Verarbeitung der Höreindrücke beteiligt ist. Die Töne der musikalischen Linie dürfen nicht als akustische Einzelreize verstanden werden. Als Phrase sind sie, rhythmisch gegliedert, eine musikalische Gestalt, die sich bei der Erarbeitung eines künstlerischen Vortrages auch als Ganzes auswirkt. In der Musik haben Vorstellung und Erlebnis des Komponisten, hervorgerufen durch den Worttext, ihren Niederschlag gefunden. Der Sprecher muß diese Vorstellungen bei der Gestaltung umsetzen. Die Absicht des Komponisten, durch die Verknüpfung von Wort und Musik eine gefühlsmäßige Vertiefung herbeizuführen, erkennen wir auch im Verlauf der melodramatischen Sprechmelodie wieder. Gesteuert von der musikalischen Linie, wird der übliche Bereich der Satzintonation

verlassen, und es entsteht eine neue emotionale Qualität. Dabei nähert sich der Tonhöhenverlauf der Sprechstimme mehr oder weniger der musikalischen Melodie. Unsere Untersuchungen bestätigen somit aus anderer Sicht die Ergebnisse von Morgon, Lafon, Prelot sowie von Siegert und Dieroff. Auch bei der auditiven Steuerung der Sprechintonation durch Musik tritt die audiophonatorische Kontrolle nicht als einfache lineare Beziehung, sondern als komplexes Phänomen der akustischen Psychophysiologie in Erscheinung.

LITERATUR

- Dieroff, H. G. u. Siegert, C.: Tonhöhenverschiebung unter Lärmbelastung. *Folia phoniat.* (Basel) 18: 247—255 (1966).
- Essen, O. v.: Grundzüge der hochdeutschen Satzintonation. (Henn, Ratingen/Düsseldorf 1956).
- Isačenko, A. V. u. Schädlich, H.-J.: Untersuchungen über die deutsche Satzintonation. *Studia grammatica*, H. 5 (Akademie-Verlag Berlin 1965).
- Morgon, A., Lafon, I. C. e. Prelot, I.-O.: Le contrôle audiophonatoire. *Folia phoniat.* (Basel) 16: 193—200 (1964).
- Reuter, F.: Der Hase und der Igel. Ein Märchen nach den Brüdern Grimm für eine Sprechstimme und Orchester, Ausgabe für Klavier. (VEB Breitkopf & Härtel, EB 4049, 1962.)
- Reuter, F.: Zur Komposition für eine Sprechstimme und Orchester (Klavier). *Wiss. Z. Univ. Halle, Ges. Sprachw.* 9: 1647—1651 (1962).
- Sedláček, K. u. Sychra, A.: Die Melodie als Faktor emotionellen Ausdrucks. *Folia phoniat.* (Basel) 15: 89—98 (1963).
- Wendler-Reuter, U. u. Wendler, J.: Zum Verhalten der Sprechstimme im Melodram. *Wiss. Z. Univ. Berlin, Ges. Sprachw.* 15: 429—438 (1966).

DISCUSSION

Kloster-Jensen:

Dr. Wendler hat zu dem ersten Punkt der gestellten Aufgabe eine Antwort gegeben, indem er es wahrscheinlich macht, daß eine begleitende Tonfolge die Satzintonation beeinflusst. In welche Weise das geschieht, wäre noch sicherer zu ersehen, wenn man dasselbe Satzstück nicht nur zu einer einzigen Melodie, sondern unter Begleitung von mehreren verschieden gestalteten Tonsequenzen sprechen ließe.

Wendler:

Ein solcher Vergleich wäre sehr interessant. Allerdings müßten dazu verschiedene, kompetente melodramatische Vertonungen des Stückes vorliegen. Eine Begleitung durch „irgendwelche“ Tonsequenzen entspräche nicht den Voraussetzungen unserer Untersuchungen; wir gingen von ganzheitlichen künstlerischen Leistungen aus.

SOME PROBLEM OF QUANTITY IN CERTAIN INDIAN LANGUAGES

E. M. WHITLEY

DISCUSSION

Lüdtke

As Mrs. Whitley has succeeded in showing, the relation between [a] and [a] is exactly on the same level as that between [a] and [a:]. This statement is, of course, not in keeping with classical phonemic theory. However, the distinction made in phonemic theory between segmental and prosodic phenomena is due to the Ancient Romans who would represent certain speech phenomena by letters, whereas they left others unrepresented. The first are called segmental, the latter prosodic. Mrs. Whitley's demonstration is just one more piece of evidence in favour of my theory which holds that phonemic theory is a consequence of the invention of alphabetic writing and subsequent development of it by Greek and Roman philosophers.

EINIGE BEMERKUNGEN ZUR ARTIKULATIONSBASIS DES RUSSISCHEN UND DEUTSCHEN

ERWIN WIEDE*

Unter Artikulationsbasis verstehen wir die für eine Sprache charakteristischen Bewegungen der Sprechorgane und die Beschaffenheit des Ansatzrohres beim Sprechen.

Die konfrontierende Darstellung der Unterschiede der Artikulationsbasen verschiedener Sprachen ist vor allem für den Fremdsprachenunterricht von Bedeutung. Für die genannten Sprachen können diese Unterschiede wie folgt skizziert werden: 1. Größere Aktivität der Lippen bei der Bildung der deutschen Vokale im Vergleich zum Russischen und die damit verbundene stärkere Labialisierung der gerundeten Vokale des Deutschen; 2. Verbindung der Artikulation der Vorderzunge mit der Lippenartikulation bei der Bildung der deutschen labialen Vokale der vorderen Reihe; 3. Neuer Einsatz bei der Bildung der deutschen Vokale im Anlaut und Fehlen des neuen Einsatzes im Russischen in dieser Position; 4. Zungenkontaktstellung beim Artikulieren deutscher Konsonanten und Velarisierung bei der Bildung der russischen nichtpalatalisierten Konsonanten; 5. Starke Aktivität des Mediodorsums bei der Artikulation der russischen palatalisierten Konsonanten; 6. Größere Aktivität der Uvula bei der Bildung der Konsonanten im Deutschen; 7. Volle Stimmhaftigkeit der russischen stimmhaften Konsonanten und Stimmtonverlust im Deutschen, besonders im Anlaut und in der Stellung nach Geräuschkonsonanten; 8. Stärkerer Expirationsstrom als im Russischen und Aspiration der deutschen $/p/$, $/t/$, $/k/$ und Fehlen der Behauchung dieser Laute im Russischen; 9. Stärkere Muskelspannung bei der Artikulation der deutschen Vokale, Sonore und der stimmlosen Konsonanten und geringere Anspannung der Muskulatur bei der Bildung der deutschen stimmhaften Konsonanten im Vergleich zum russischen Lautbestand; 10. Fließende Artikulationsbewegungen (Koartikulation) und die damit verbundene stark ausgeprägte Assimilation im Russischen und stärkeres Aushalten der Artikulationsbewegungen im Deutschen.

* DDR, Güstrow, Pädagogisches Institut.

Die Artikulationsbasis einer Sprache steht in enger Wechselbeziehung zum phonologischen System dieser Sprache. Die Abhängigkeitsverhältnisse sind in der Diachronie und Synchronie jedoch unterschiedlich.

In der Diachronie beeinflußt die Artikulationsbasis die Evolution des phonologischen Systems der Sprache. So konnten z. B. die palatalisierten Konsonanten im Russischen zu Phonemen werden, weil sie durch die Artikulationsbasis des Gemeinslavischen dazu „vorbereitet“ worden sind. Hier wurden die Konsonanten in der Position vor Vokalen der vorderen Reihe palatalisiert gesprochen. Die palatalisierten Konsonanten entwickeln sich somit zunächst als Allophone der nichtpalatalisierten Konsonantphoneme. Die nichtpalatalisierte bzw. palatalisierte Aussprache der Konsonantphoneme war in dieser Periode eines der Merkmale der Artikulationsbasis des Gemeinslavischen. Auch die Entwicklung der deutschen labialen Vokale der vorderen Reihe zu Phonemen könnte hier als Beispiel angeführt werden.

In der Synchronie dagegen beeinflußt das phonologische System einer Sprache die Spezifik der Aussprache. So kann die stark ausgeprägte Bewegung des Mediodorsums beim Bilden der palatalisierten und die Velarisierung bei der Artikulation der nichtpalatalisierten russischen Konsonanten nur durch das Vorhandensein der Opposition palatalisierter/nichtpalatalisierter Konsonantphoneme erklärt werden.

Die Besonderheiten der Artikulationsbasis einer Sprache stehen auch mit dem morphologischen System dieser Sprache in Beziehung. Das zeigt sich z. B. beim neuen Einsatz im Deutschen. Durch den neuen Einsatz können hier Morpheme und Wortformen voneinander unterschieden werden. Vgl. z. B. Wörter und Wortverbindungen wie:

mitteilen und miteilen,
verreisen und vereisen,
am Mast und am Ast,
der Riegel und der Igel.

ON PHONOLOGICAL AMBIGUITY

G. WIENOLD*

1. Within the phonological component of the theory of generative grammars problems of phonological ambiguity arise. A phonetic segment will be called phonologically ambiguous, if it cannot be decided on what the underlying abstract phonological (systematic phonemic) representation of this segment should be. This situation has to be kept strictly apart from that where a phonetic segment in some occurrences will be assigned one abstract representation and in some other occurrences another. In this case, there will be a unique abstract representation of each occurrence of the phonetic segment, and by applying the phonological rules the phonetic segment will be generated correctly from the postulated abstract representation. In the case of phonologically ambiguous segments there will be more than one abstract representation of the segment from which it can be generated correctly by applying the phonological rules. The reduced vowel in English provides a good example. The vowel of the last syllable of words like *constitution*, *relaxation*, *division* allows for more than one abstract representation, all of which would yield the correct and expected results.¹

2. Before continuing with the discussion of this problem I want to comment briefly on the notion of ambiguity. It has been among the particular concerns of the theory and practice of generative grammar to provide descriptions and explanations of syntactic and semantic ambiguities. A given sentence is called ambiguous if it admits of more than one interpretation. Two or more sentences may be identical on the surface but differ in meaning because they are derived from different underlying phrase-markers. This is called *syntactic ambiguity*. Consider, e.g., the sentence:

The police were ordered to stop drinking after midnight.

Two or more sentences may be identical in surface structure but differ in meaning because more than one lexical reading may be associated with one or more than one

* Englisches Seminar der Universität Münster/Westf.

¹ There seems to be nothing in print on this matter. I can only refer the reader to Chomsky's remark that the occurrence of the reduced vowel "can largely be predicted by extension of the transformational cycle" (Noam Chomsky, "Topics in the Theory of Generative Grammar", in Thomas A. Sebeok [Ed.], *Current Trends in Linguistics*, vol. III [The Hague, 1966], p. 56).

of the terminal elements of their underlying phrase markers. This is called *semantic ambiguity*. Consider, e.g., the sentence:

John prefers light beer.

In both cases of syntactic and semantic ambiguities, the ambiguity will be resolved, that is, with each meaning of an ambiguous phrase there will be associated a unique syntactic description and a unique semantic interpretation.

Thus, phonological ambiguity as described here differs from syntactic and semantic ambiguities in that it is not resolved.

3. Consider the case of some Gothic vowels. Ronald E. Buckalew² presented a generative morphology of Gothic where each occurrence of the segment [ɛ] is generated from an underlying *i*, *e*, or from the sequence *ai*, each occurrence of the segment [ɔ] is generated from an underlying *u*, *o*, or the sequence *au*. I use letter notation as an abbreviative device instead of feature matrices. Here, I will discuss the derivations of [ɛ] from *i* or *ai*, and of [ɔ] from *u* or *au*, only. As time does not permit I cannot relate these statements to traditional phonemic analyses of Gothic.³ In the analysis referred to, [ɛ] and [ɔ] usually have as their abstract representations *i* and *u*, respectively, when they occur before *r*, *x*, *xʷ*, in other cases they will be represented as *ai* or *au*. This analysis is based on alternations like *bai*—*bajops*, *mawi*—*maujos*. Apart from some marginal exceptions, which I shall disregard, there are some systematic ones. They occur, for instance, in the past tense of the irregular (strong) verbs.

Consider the following verb forms. The table gives in the first column the graphemic notation in a modern transliteration, in the second column the assumed phonetic reading, in the third column, their abstract representations.

1 sg. past	gataih	gatex	gataix
1 pl. past	gataihum	gatexum	gatixum
1 sg. past	tauh	tɔx	taux
1 pl. past	tauhum	tɔxum	tuxum

As stated above, there are rules that convert *i*'s and *u*'s before *x* into [ɛ]'s and [ɔ]'s of the phonetic representation. There are, as far the problem under discussion is concerned, no environments specified for the rules that apply to *ai* and *au*. They will be converted into [ɛ]'s and [ɔ]'s of the phonetic representation in all instances. This means that the abstract representations notated as *gataix* and *taux* could, as well, be notated *gatix* and *tux*, and vice versa. The correct phonetic representation would be generated all the same. Now, in this case, there are rather strong motivations to adopt the abstract representation as given in the table above. By these motivations the ambiguity is resolved. It is known that verbs belonging to the same classes but

² Ronald E. Buckalew, *A Generative Grammar of Gothic Morphology*, Diss. Urbana, Illinois, 1964.

³ For bibliographical references see my paper, "The Pre-Gothic Monophthongizations and Wulfila's Graphemic System", (to appear).

not containing *x* (or *r* or *xʷ*) as stem final consonant display the same alternations *ai* ~ *i*, *au* ~ *u* in the past tense.

1 sg. past	graip	grɛp	graip
1 pl. past	gripum	gripum	gripum
1 sg. past	baup	bɔp	baup
1 pl. past	budum	budum	budum. ⁴

This systematic correspondence justifies the representation of <gataih> as *gataix* and of <tauh> as *taux*, the representation of the plural forms <gataihum> and <tauhum> as *gatixum* and *tuxum*.

4. One will notice that the abstract representations are, by and large, the forms one would reconstruct for pre-stages of Gothic. Thus, the analysis would support the claim made by generative phonology that the abstract representations change much more slowly than the surface phonetic representations.⁵

There are other cases that cannot be resolved this way. The past tense verb forms show that the segments *ɛ* and *ɔ* before *r*, *x*, and *xʷ* may have *i* and *u*, or *ai* and *au* as their abstract representations. In cases like the noun <baur> [bɔr], the adjective <gaurs> [gɔrs] 'sad' there is no way to decide whether *au* or *u* should be the abstract representation, except by appealing to etymology which is not allowed. The same is true of the reduplicating syllable occurring in the past tense of a certain class of verbs and consisting of the initial consonant (or consonant cluster in some cases) and the vowel [ɛ] as in *hopan*—*hɔihop*, *letan*—*lɔihlot*. As far as the rules have been stated, [ɛ] in *lɔihlot* should have as its abstract representation *ai*, [ɛ] in *hɔihop* *i* or *ai*, without a possibility of deciding between both solutions. Historically, in both cases, [ɛ] is to be derived from **i*. The representations *ai* and *i* for [ɛ] in *hɔihop* are arbitrary, [ɛ] is phonologically ambiguous. Historically, this means that Pre-Gothic *ai* and *i* have merged not only on the surface but also on the abstract phonological representation.

5. Two contrasting conclusions seem possible. If we admit descriptions containing phonological ambiguities, this would allow in the Gothic example to analyze, among other things, paradigms like <gataih>—<gataihum> in the same way as <graip>—<gripum>, and by this we would retain a maximum of generalization on the behavior of verb classes.⁶ This description of Gothic, however, would contain a number of forms unique abstract representations of which could not be stated, the grammar would contain two (or more) alternative abstract representations on the status of

⁴ Phonetic and phonological representation as *budum* would be acceptable as well.

⁵ Cf. Morris Halle, "On the Rule of Simplicity in Linguistic Descriptions", in Roman Jakobson [Ed.], *Structure of Language and its Mathematical Aspects* (Providence, 1961), pp. 89—94; idem, "Phonology in Generative Grammar", in Jerry A. Fodor and Jerrold J. Katz (Eds.), *The Structure of Language: Readings in the Philosophy of Language* (Englewood Cliffs, New Jersey, 1964), pp. 347, 351f.

⁶ On generalization cf. Chomsky, l. c., pp. 49 ff. (with literature).

which only arbitrary statements could be made. Alternatively, one would not allow for generalizations beyond a point where phonological ambiguities arise in consequence of such generalizations. Such a grammar of Gothic would generate all segments [ɛ] and [ɔ] uniquely. (This abstract representation would be, by the way, closer to the phonological analysis of Gothic which is now more and more accepted.) Some subparts of the grammar, e.g. verb morphology, would contain less generalization. And in this part, the grammar 2 would also deviate from traditional textbooks, as they present the same historical analysis that is preserved in the alternative grammar 1 with a maximum of generalization.

AUS DEN KINORÖNTGENOGRAPHISCHEN UNTERSUCHUNGEN DER POLNISCHEN VERSCHLUSSLAUTE

BOŻENA WIERZCHOWSKA

Wie aus den kinoröntgenographischen Untersuchungen hervorgeht, werden die Realisationen der polnischen Phoneme *p, b, t, d, k, g* und der entsprechenden, weichen Phoneme *p', b', t', d', k', g'*, die normalerweise als Verschlusslaute aufgefaßt werden, nicht immer mit einem vollen Verschuß der entsprechenden Sprechorgane, d. h. der Lippen, der Zungenspitze usw. gebildet. Meistens ist der Verschuß bei ruhigem, gepflegtem, in einem langsamen Tempo geführten Sprechen zu beobachten und er dauert ca 0,06—0,15 Sek. Bei einer Filmgeschwindigkeit von ca 50 Bilder pro Sekunde sind die Verschlüsse auf 3 bis 7 nacheinander folgenden Bildern zu beobachten. Bei schnellem Sprechen sind die Verschlüsse schwächer und manchmal sind sie sogar nicht komplett; vom Röntgenfilm werden sie überhaupt nicht erfaßt. Solche schwächere, unkomplette Verschlüsse sind häufiger bei den Realisationen der labialen, postalatalen und velaren Verschußphonemen als beim Realisieren der dentalen Phoneme.

Diese erwähnten Abweichungen bei der Realisierung von polnischen Verschußphonemen können die Benutzer der Sprache unbeachtet lassen, sogar auch dann, wenn diese Benutzer phonetische Ausbildung besitzen. Diese Abweichungen sind meistens erst durch die phonetische Analyse der kinoröntgenographischen Filme zu erfassen.

Bei sehr schnellem Sprechen und mit nicht immer gepflegter Aussprache kommt es manchmal, besonders bei geläufigen und öfters gebrauchten Wörtern und Redewendungen, zu einer Assibilation bei Realisierungen der Verschußphoneme — z. B. das Phonem *g* ist in solchen Formen wie: *voy le* (in der gepflegten Aussprache *v ogule* — „meistens“) als Engellaut γ realisiert. Dieser γ -Laut wird in Realisierungen des *g*-Phonems schneller artikuliert als in Realisierungen eines *x*-Phonems.

Bei der Realisierung der Verschußphoneme scheint der wichtigste Moment der Artikulation die schnelle Bewegung der entsprechenden Artikulationsorgane (d. h. die schnelle Bewegung der Lippen, der Zungenspitze usw.) zu sein; dabei ist ein voller Verschuß sowie auch eine Explosion weniger wichtig. Diese schnelle Bewegung kann nur eine rasche Annäherung der Lippen, eine schnelle Annäherung der Zungenspitze an die vorderen Schneidezähne oder des Zungenrückens an den Gaumen usw. sein. Eine schnelle Bewegung kann auch eine Implosion sein — so wie bei den Implosivlauten, die in der polnischen Sprache in den Interjektionen *hop! hap!* vorkommen.

Dies kann auch eine laterale Explosion sein: in den Lautgruppen tl, dl, kl (poln. tlen „Sauerstoff“, dla „für“, klasa „Klasse“), oder eine nasale Explosion in den Lautgruppen tn, dn, pm, bm (poln. tnę „ich schneide“, rupmy — „machen wir“).

Werden die Laute t, d in der Nachbarschaft von Vokalen mit Hilfe einer ähnlichen Lage der Zungenmasse artikuliert, wie z. B. in tata („Vati“), da („er wird geben“), kommt es öfter vor, daß der ganze Unterkiefer von dem Oberkiefer abspringt. Jede von den hier genannten schnellen Bewegungen der Sprechorgane verursacht die Entstehung eines kurzen akustischen Impulses, der als ein Explosivlaut wahrgenommen wird. Zur Erzeugung eines solchen akustischen Impulses genügt eine schnelle Bewegung der Sprechorgane, die eine rasche Veränderung des akustischen Druckes verursacht. Die Wahrnehmung der so artikulierten Sprachlaute ist der Perzeption von künstlich gekürzten Engelaute (Frikativlaute) ähnlich — die so gekürzten Engelaute werden als Verschußlaute perzipiert. Eine schnelle Bewegung der entsprechenden Sprechorgane (also genetisch gesehen), sowie eine schnelle Veränderung des akustischen Ablaufes, rufen einen ähnlichen auditorischen Effekt hervor.

Das menschliche Nervensystem verarbeitet leichter Informationen, welche die Bewegungen der Muskeln verursachen, als Informationen, die durch den Kontakt der Zunge mit der wenig innervierten Schleimhaut der Mundhöhle erzeugt werden.

LAUT, BEDEUTUNG UND SPRACHGRUNDEINHEIT

JÓZEF WIERZCHOWSKI

In der Struktur der Sprache betont man Elemente, die als sogenannte Einheiten aufgefaßt werden. Diese Elemente werden nicht einheitlich in verschiedenen linguistischen Forschungsrichtungen betrachtet, im Zusammenhang mit zwei verschiedenen, extremen Forschungsrichtungen in der Linguistik werden sie meistens als Elemente der Form oder als Elemente des Inhalts dargestellt; wenigstens der Ausgangspunkt ist Form oder Bedeutung.

Produktiv scheint die Analyse der Sprachstruktur zu sein: diese versucht die Sprachgrundeinheit zu definieren, das heißt, daß sie die Form und Bedeutung einheitlich und gleichzeitig betrachtet. Eine so durchgeführte Analyse soll also als Ergänzung zur bekannten Analyse, die meistens von der Form zur Bedeutung, oder umgekehrt, von der Bedeutung zur Form ausgehend verwendet wird.

Ferdinand de Saussure postulierte das Wort als Grundeinheit der Sprache. Die Methode der einheitlichen und gleichzeitigen Auffassung von Form und Bedeutung stammt von Kruszewski und Baudouin de Courtenay und ist mit der Prager Schule eng verbunden. Die Methode des Kommutationstestes führt gleichzeitig zu Aufstellung des Phoneminventars und des Inventars der Wortformen natürlicher Sprachen. Diese Analyse berücksichtigt adäquat sowohl die Form als auch die Bedeutung. Obwohl sie die Form und Bedeutung adäquat berücksichtigt, bietet sie dennoch nicht eine linguistisch ausführliche Darstellung. Die so aufgefaßten Elemente der Struktur der Sprache, die Wortformen, die wir auch als Einheiten bezeichnen können, berücksichtigen die Bedeutung nur distinktiv; sie sind meistens aber polysemantisch. Polysemantische Elemente der Sprache können nicht als Grundeinheiten aufgefaßt werden. Die Grundeinheit der Sprache muß bezüglich der Bedeutung einheitlich sein. Wir sind demnach also gezwungen, polysemantische Wortformen in einzelnen Bedeutungen einzuteilen. Eine polysemantische Wortform enthält also mehrere phonemisch-homophonische Wörter. Die Bedeutungen dieser Wörter müssen aber einheitlich aufgefaßt werden. Erst solche Elemente, die einheitliche Bedeutung haben, können als Grundeinheiten der Sprache angesehen werden.

Die Durchführung einer Teilung der Wortformen in Wörter (Grundeinheiten der Sprache) sollte die Aufgabe der Semantik sein. Eine Semantik, die eine solche Durchführung erreichen könnte, muß also ihren Forschungsschwerpunkt von der Betrachtung

tung der Beziehungen zwischen Form und dem Bezeichneten auf die Relationen der bedeutenden Sprachformen verlagern. Demnach müssen wir also die ganze Fächerung der Wortform analysieren. Die Grundlage zur Durchführung der Einteilung von Wortformen in einzelne Bedeutungen, d. h. in einzelne Wörter, ist die Verschiedenheit in der Fächerung in Form und Bedeutung. Z. B. in der Wortform kritisch sind mindestens zwei verschiedene Wörter enthalten. Erstens: das Wort *kritisch* in der Relation zur *kritischen Lage*, zur *kritischen Temperatur*. Zweitens: das Wort *kritisch* in der Relation zur *kritischen Äußerung*, *kritisieren*, *Kritiker*. Das erste Wort *kritisch* befindet sich nur in der formalen Relation zu den Formen *Kritiker*, *Kritik*, *kritisieren*, *kritische Äußerung*. Die Beobachtung der ganzen Fächerung der Formen gibt uns eine Methode, welche die Durchführung der Teilung von Wortformen in Wörter ermöglicht. Diese Methode erlaubt uns den lexikalischen Vorrat einheitlich einzuteilen.

Die hier erwähnte Methode, die gleichzeitig und einheitlich Form und Inhalt berücksichtigt, ist für die Beschreibung der Struktur der Sprache von Bedeutung, d. h. für eine solche Beschreibung, in der das Problem der Bedeutung einen speziellen Wert hat. Sie führt auch zur ausführlicheren Berücksichtigung der Bedeutung bei der Beschreibung der Struktur der Sprache. Diese Methode gibt uns Forschungsmöglichkeiten nicht nur auf der Ebene der Semantik der Lexikologie und Morphologie, sondern auch in der Phonologie und Phonetik. Besonders fruchtbar ist sie bei Forschungen der Realisierung von Phonemen. Sie ermöglicht nicht nur die Beobachtung phonemischer Realisationen in bezug auf das Phoneminventar, die Nachbarschaft von Phonemen und die Wortformen, sondern auch die Beobachtung phonemischer Realisationen in bezug zur Bedeutung. Eine so durchgeführte Analyse erlaubt statistisch festzustellen, daß es in den Phonemrealisationen Verschiedenheiten gibt, die wir mit den Bedeutungen der Formen verbinden müssen. Diese Methode ist für die diachronische Analyse der Sprache bedeutungsvoll. Schwierigkeiten bei der Bestimmung sprachhistorischer Gesetze ergeben sich oft deshalb, weil wir als Ausgangspunkt der Forschung nicht Wörter, sondern polysemantische Elemente der Sprache — Wortformen — betrachtet haben.

DISCUSSION

Bergsveinsson:

Der Vortragende betont den Unterschied zwischen „Wort“ und „Wortform“. Beispiel: „Kritisch“, 1. als Lage, Krankheit, 2. als Stellungnahme. Die Wortform ist hier polysemantisch. Das Wort als Grundeinheit darf es nicht sein, hat also entweder die erste oder zweite Bedeutung. Ich mache darauf aufmerksam, daß „Wort“ an sich in der Fachliteratur mehrdeutig ist, kann daher nicht als Grundeinheit gelten. Die Termini „Wort“ und „Wortform“ sind zu ungenau, um die betreffenden Begriffe in der empirischen Wissenschaft anzuwenden.

DIE BEURTEILUNG MUNDARTLICHER AUSSPRACHEN DURCH MUNDARTSPRECHER IM DEUTSCHEN

PETER WIESINGER*

Dem Generalthema dieses Kongresses „Akustisches Sprachsignal und seine Perception“ entsprechend, mag es auch einmal interessant sein zu erfahren, welche Vorstellungen die Sprecher selbst von den von ihnen verwendeten Lauten haben und nach welchen Maßstäben sie diese beurteilen. Gleichzeitig ergibt sich die Frage, ob sich die Lautvorstellungen der Sprecher mit der objektiven Beurteilung des Phonetikers und Linguisten decken oder ob hier Unterschiede bestehen. Bei möglichen Abweichungen in der Auffassung ist schließlich die Frage zu klären, wieso es zu diesen kommt. Wir wollen dieses Problem im folgenden an ausgewählten Beispielen aus dem Vokalismus deutscher Mundarten in Grundzügen zu erläutern versuchen.

Bei der Untersuchung gesprochener Sprache, speziell einer Mundart, ist der Linguist bestrebt, zunächst das Lautinventar phonetisch festzuhalten und zu sichten und dann daraus das Phoneminventar abzuleiten, d. h. aus der Fülle unterschiedlich artikulierter Laute jene Lauttypen festzustellen, denen bedeutungsunterscheidende Funktion zukommt und die damit Phoneme sind. Das aus dem Lautinventar gewonnene Phoneminventar kann nicht nur für den Linguisten Bedeutung besitzen, sondern muß auch dem Sprecher selbst gegenwärtig sein, da es ja die Phoneme sind, die als kleinste Bausteine der Wörter die Verständlichkeit der sprachlichen Äußerungen gewährleisten und für das Funktionieren der Sprache als Kommunikationsmittel der Gemeinschaft der Sprachträger, kurz der Sprachgemeinschaft, verantwortlich sind. Es wäre freilich falsch, aus dieser Tatsache auf die Kenntnis des Phoneminventars als abstraktes, analytisches System im Bewußtsein der Sprecher schließen zu wollen. Aber ein Gefühl für die Unterscheidungskraft verschiedener Laute und damit ein unbewußtes Empfinden für den Phonemcharakter von Lauten ist den Sprechern durchaus eigen. Unmittelbar läßt sich das überprüfen, wenn man einen Mundartsprecher zur Beurteilung der Reimbarkeit oder Nichtreimbarkeit von Wörtern auffordert, eine Fähigkeit, die durch jahrhundertelange Tradition nicht nur anerzogen ist, sondern geradezu schon angeboren zu sein scheint, und die stets das gewünschte Ergebnis bringt. Eine indirekte Möglichkeit zur Überprüfung bietet dem Linguisten die diachrone Methode. Es gibt keine deutsche Mundart, in der nicht

* Forschungsinstitut für dt. Sprache, Deutscher Sprachatlas, Marburg/Lahn.

ursprünglich getrennte Phoneme durch lautgesetzlichen Wandel gleich geworden und damit zusammengefallen wären, was Homonyme oder besser Homophone zur Folge hatte. Es ist nun bezeichnend, daß der lautgesetzliche Wandel in Einzelbeispielen gelegentlich durchbrochen worden ist, wenn die dadurch entstehenden Homophone zu Bedeutungsmissverständnissen geführt hätten. So wären z. B. in der rheinfränkischen Vorderpfalz nach der Apokope von -e durch den Zusammenfall von mhd. *g* und mhd. *ch* im Auslaut die Wörter „Lauch“ und „Lauge“ als *lɛ:ç* homophon geworden, was in der Küche wegen der ständigen Verwendung der Gemüse- und Gewürzpflanze und des Reinigungsmittels zu unliebsamen Verwechslungen geführt hätte. Aus diesem Grund ist daher in einem der beiden Wörter die lautgesetzliche Palatalisierung des Vokals unterblieben, und werden heute lautgesetzliches *lɛ:ç* ‚Lauge‘ und nicht lautgesetzliches *la:x* ‚Lauch‘ unterschieden, worin sich ebenfalls das Empfinden für phonematische Unterscheidungen spiegelt.

Das Vorhandensein eines „Phonemgefühls“ befähigt aber den Sprecher noch keineswegs zur Beurteilung der von ihm gesprochenen Laute. Dazu bedarf es der Vergleichsmöglichkeiten. Sie bestehen für einen in abgelegenen Gegenden lebenden Mundartsprecher, der weder die Möglichkeit hatte, Lesen und Schreiben zu lernen, noch jüngst durch den Rundfunk eine höhere Sprachschicht vermittelt zu bekommen, höchstens in abweichenden Nachbarmundarten seiner näheren Umgebung. Was da an lautlichen Unterschieden auffällt, wird gerne registriert und oftmals in Spottsprüchen lustig gemacht. Man kann auf solche Weise zwar immer erfahren, was der Nachbar anders spricht und wie er es spricht, doch reichen diese Beobachtungen nicht zur Beurteilung der Laute aus. Diese setzt eine gewisse Schulung in sprachlichen Dingen voraus, zumindest das Lesen und Schreiben und damit die Kenntnis der Schrift- oder Hochsprache. Sie wird in ländlichen Gegenden, wo die Kinder im Gegensatz zu vielen Städten noch mit einer Mundart aufwachsen, geradezu als eine andere Sprache erlernt, wobei in der Schulung des Schreibens der Lautwert der Schriftzeichen bewußt wird. Gleichzeitig fordert die Kenntnis der Hochsprache als einer anderen, höheren Sprachschicht zum Vergleich mit der eigenen Mundart heraus. Unter solchen Voraussetzungen ist es möglich, die Auffassung des Sprechers über die von ihm verwendeten mundartlichen Laute zu erfahren.

Dabei besteht keine Schwierigkeit in der Beurteilung oraler, palataler oder velarer Vokale wie *i*, *y*, *e*, *ø*, *u*, *o*, *a*. Sie haben ihre Entsprechungen in der Hochsprache und werden als solche registriert, auch wenn die mundartlichen Aussprachen von der Hochsprache abweichen, z. B. *si:* statt *se:* ‚See‘, *ru:t* statt *ro:t* ‚rot‘ u.s.w. Besteht aber eine Opposition zwischen zwei Vokalen desselben Typus z. B. *lu:s* ‚Lau‘ gegenüber *lu:ɔ* ‚lose‘, so werden zwar beide Vokale als *u* bezeichnet, doch seien sie nicht gleich sondern „anders“, was wieder das „Phonemgefühl“ des Sprechers bestätigt.

Schwieriger wird es bereits bei nasalen Vokalen, deren „durch die Nase gesprochene“ Klangfarbe nur dann erkannt wird, wenn sie sich noch deutlich den obengenannten Grundtypen zuordnen läßt. Ist dies unmittelbar nicht mehr möglich und fehlt das hochsprachliche Äquivalent, wie z. B. bei *m̃:* ‚Mann‘ mit sehr offenem *o*, dann

greift man gerne zum hochsprachlichen Wort und vermutet im mundartlichen ein „durch die Nase gesprochenes“ *a*. Fragt man dann aber den Sprecher, wie er sein *k̃ä:* ‚kein‘ beurteilt, so gerät er freilich in Verlegenheit und muß *a* diesem Beispiel zuordnen, während er das andere nun zu *o* stellt. Nennt man schließlich noch *t̃ō:* ‚Ton‘, so kann man bestenfalls noch hören, daß es ein anderes *o* als in *m̃o:* ist, was zwar wieder das Gefühl für phonologische Unterscheidbarkeit beweist, aber schon die Fähigkeiten des Sprechers überfordert.

Ein besonderes Problem stellen die in den deutschen Mundarten sehr verbreiteten einst velaren Zentralvokale dar, die als *ü* und *ö* noch mehr an die velaren *u* und *o* und als *ü* und *ö* schon mehr an die palatalen *y* und *ø* anklängen. Hier werden erstere von den Sprechern stets als *u* und *o* aufgefaßt, während die Beurteilung der letzteren geteilt ist. Als Beispiel dafür sei die osthessische Mundart um Hersfeld—Rotenburg gebracht, wo mhd. *û* stets als stark zentralisiertes *ü:* gesprochen wird, z. B. *hü:d* ‚Haut‘. Während dieses um Hersfeld trotz seines *y*-Klanges von den Sprechern als *u* klassifiziert wird, wird derselbe Laut um Rotenburg richtig als *y* verstanden. Die Lösung dieses Rätsels ist im phonologischen System beider Mundarten zu suchen: Während die Hersfelder Mundart kein langes Phonem */u:/*, sondern nur das stark zentralisierte Phonem */ü/* kennt, das in nächster Opposition zu palatalem */i:/* für mhd. *î + û*, z. B. *hi:d* ‚Häute‘, und zu velarem */o:/* für mhd. *uo*, *ô* und Dehnungs-*ō*, z. B. *ho:d* ‚Hut‘, steht, verfügt die Rotenburger Mundart neben dem Phonem */ü:/* für mhd. *û* auch noch über ein Phonem */u:/* für mhd. *uo* z. B. *hu:d* ‚Hut‘, deren nächste Oppositionen wieder die Phoneme */i:/* für mhd. *î + û* und *ie + üe* und */o:/* für mhd. *ô* und Dehnungs-*ō* bilden. Das dem Hersfelder System fehlende *u:* ermöglicht also im Gegensatz zum Rotenburger System die Wertung von *ü:* als *u:*.

Schließlich verbleibt noch die Bewertung der Diphthonge, deren Beurteilung als *ai*, *oi*, *au*, *iə*, *ya*, *ua*, *ea*, *øa*, *oa* keine Schwierigkeiten bereitet. Während *ei*, *øy*, *ou* weitgehendst noch erkannt werden, versteht man diphthongoides *ei*, *øy*, *ou* allgemein monophthongisch, weil einerseits der Klangeindruck der langen *i*. Komponenten überwiegt und andererseits keine Opposition zu entsprechenden langen Monophthongen *e:*, *ø:*, *o:* besteht. Das ist besonders in weiten Teilen des Nordniederdeutschen der Fall, wo die nächsten monophthongischen Oppositionen */i:/*—*/y:/*—*/u:/* und sehr offenes */ɛ:/*—*/æ:/*—*/ɔ:/* und die einzigen diphthongischen */ai/*—*/oy/*—*/au/* sind. Gerät hier selbst der Linguist in Zweifel, ob er solche Aussprachen dem Monophthong- oder Diphthongsystem zuordnen soll, so kann er bezüglich der überoffenen Monophthonge *æ:* für mhd. *î + û* und *v:* für mhd. *û*, z. B. *æ:s* ‚Eis‘, *v:s* ‚aus‘ im Mittelbairischen um Wien und im Nordbairischen um Viechtach—Furth im Bayrischen Wald nicht im Zweifel sein. Dennoch halten die mittelbairischen Sprecher diese Monophthonge im Gegensatz zu den nordbairischen für Diphthonge *ai*, *au*. Auch hier verhelfen die phonologischen Systeme zur Lösung: Da es um Wien keinerlei steigende Diphthongphoneme gibt, dafür aber diese überoffenen Monophthonge zu weniger offenen *ɛ:—ɔ:* und geschlossenen *e:—o:* in Opposition stehen, also der *E—O*-Bereich sehr stark besetzt ist, der Diphthongbereich aber gar nicht, verleitet der Vergleich mit der

Schriftsprache zur falschen Zuordnung. Anders im Bayrischen Wald, wo es die Diphthongphoneme /ei/—/ou/ für mhd. ie+üe — uo und /ɛi/—/ɔu/ für mhd. ê + ô — ô gibt, die eine falsche Beurteilung von /æ:/—/ɐ:/ nicht zulassen.¹ Der Linguist hat natürlich in beiden Fällen eine monophthongische Zuordnung zu vollziehen.

Fassen wir zusammen, so ergibt sich, daß 1. den Mundartsprechern ein „Phonemgefühl“ für die Bedeutungsunterscheidung der Laute eigen ist; daß 2. zur bewußten Beurteilung von Lauten die Kenntnis der Schriftsprache (Hochsprache) als Vergleichsobjekt notwendig ist; daß 3. die richtige Zuordnung der in Hochsprache und Mundart in gleicher Weise vertretenen Laute keine Schwierigkeiten bereitet; und daß 4. spezifisch mundartliche, der Hochsprache fremde Laute mit den nächst verwandten hochsprachlichen Äquivalenten in Verbindung gebracht werden, wobei auf Grund der phonologischen Oppositionen auch falsche phonetische Zuordnungen getroffen werden können.

DISCUSSION

Beyer:

On ne peut pas toujours se fier au „sentiment“ phonologique de l'informateur (un exemple précis emprunté à l'alsacien est cité). D'autre part, de nombreux linguistes sont d'accord sur ce point, l'enquête sur le „terrain“ ne peut être que phonétique. A propos des collisions homonymiques, il faut préciser qu'en germanique elles ont souvent été évitées grâce au procédé de la composition (cf. alsacien Sais „scie“, Maisais „faux“, Fischotter „loutre“, Kreuzotter „vipère“, etc.

Heike:

Der Terminus „zentralisierter Velarvokal“ enthält in synchronischen Betrachtungen eine unzulässige Vermischung mit diachronischen Aussagen. Es gibt entweder Zentralvokale oder Velarvokale oder Zentralvokale, die Velarvokalen eines älteren Zustandes des Systems entsprechen.

Martens:

1. Bei der Zuordnung zu einem bestimmten Phonemwert (durch den Mundartspreeher) kann — bei Hochdeutsch geführtem Gespräch — außer der Lautnorm in der Vorstellung auch noch die Aussprache des Explorators als Leitbild sich auswirken. Das sollte berücksichtigt werden.

2. „Die Phoneme müssen dem Sprecher bewußt sein“ — „Er muß ein Gefühl für den Phonemwert von Lauten haben“ sind u. U. mißdeutbare Formulierungen; vielleicht ist es glücklicher zu sagen Gefühl für die *Relationen der Phoneme zueinander*.

¹ Aus Gründen der besseren Verständlichkeit geben wir hier in beiden Fällen lange Quantitäten an, obwohl im Mittelbairischen die Quantität phonologisch irrelevant ist.

DARSTELLUNG DES SPRECHVERHALTENS ALS STATISTISCHE TONHÖHEN - UND FORMANT- VERTEILUNG MITTELS LANGZEITANALYSE

F. WINCKEL

Wenn man davon ausgeht, daß die Lautproduktion der Sprache einen stochastischen Prozeß darstellt, so muß man erwarten, daß man in einer statistischen Auswertung der Lautparameter Merkmale der Sprecher von Sprachgemeinschaften ermitteln kann. Die Auftragung der Amplituden-Häufigkeitsverteilung über dem Sprachfrequenzband, ermittelt aus dem Sprechvorgang von mehreren Minuten, müßte Merkmale einer Sprachgruppe erkennen lassen.

Jedoch nur bei gedrucktem Text werden die statistischen Stichproben an verschiedenen Abschnitten des Textes den gleichen Mittelwertverlauf ergeben (ergodisches Verhalten), während in der Lautsprache die prosodemischen Merkmale die ergodische Lautstruktur verdecken, wobei diese wiederum von emotionalen Faktoren abhängen.

Wenn es gelingt, identische normale Verteilungskurven von Sprechern einer Gruppe zu ermitteln, dann müssen die Abweichungen von diesen Kurven die prosodemischen Merkmale als quantitative Größen aufweisen. Die Möglichkeit läßt sich anhand der hier gezeigten Histogramme nachweisen.

Hierzu wurde ein Gerät entwickelt, das in einer Terzfilteranalyse die Häufigkeit der Amplituden, die einen mittleren Lautpegel — hier z. B. 30 dB — unter Vollaussteuerung überschreiten, mittels eines Zählers anzeigt.

Die relative Terzpegelhäufigkeit ergibt sich, indem die Impulszahl Z auf die Mittenfrequenz des jeweiligen Terzfilters f_m und die Gesamtsprechdauer t_{ges} bezogen wird.

Terzpegelhäufigkeit

$$H = \frac{Z \cdot 100}{f_m \cdot t_{ges}} \quad (\%) \quad (1)$$

Es ergibt sich im allgemeinen eine mit der Frequenz abnehmende Häufigkeitsverteilung (Abb. 1a). In einer anderen Darstellung wird der Mittelwert der Impulszahl über der Gesamtmeßzeit als Funktion der Terzfilter-Mittenfrequenz gezeigt (Abb. 1b).

Mittlere Zählrate

$$Z_m = \frac{Z}{t_{ges}} \quad (2)$$

In diesem Falle kommen die höheren Frequenzgebiete besser zur Geltung, wodurch das Übergangsgebiet vom Vokal- zum Konsonantbereich genauer untersucht werden

kann. Im Fall (1) tritt der tieffrequente Bereich besser hervor, weshalb er über die Tonhöhenverteilung besser Auskunft gibt. Die Häufigkeitskurven lassen oft einen Einbruch im Bereich von 200 Hz (männliche Stimme) erkennen, was eine

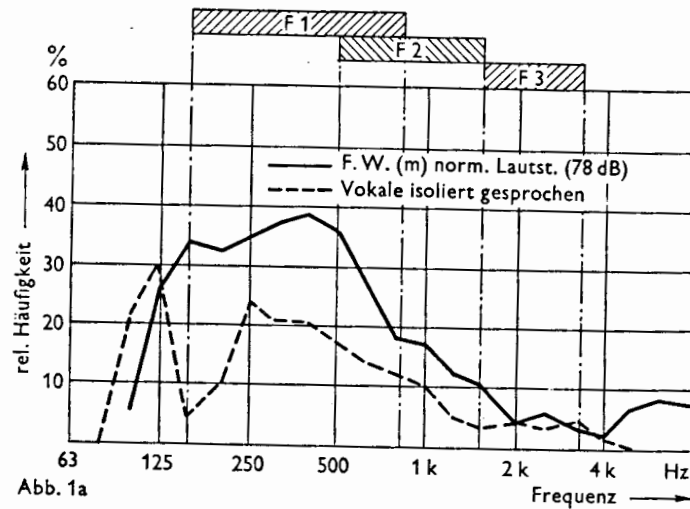


Abb. 1a

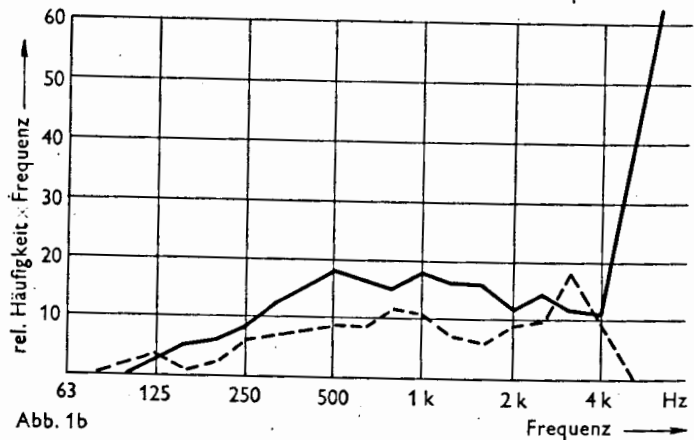


Abb. 1b

Abtrennung des Grundtonbereichs vom Formantbereich bedeutet. Zur Bestimmung der Tonhöhen-Häufigkeitsverteilung wurde in einer getrennten Messung eine feinere Filterteilung im Bereich unterhalb 200 Hz vorgenommen (Abb. 2b). Die Amplitudenverteilung unabhängig von der Frequenz wurde schon früher als eine Gaußverteilung festgestellt.

Abb. 3 zeigt das Sprechverhalten von zwei männlichen und einem weiblichen Sprecher beim Lesen von belanglosem Text über 3 Min. Die Abweichung von der ergodischen Struktur ist vor allem in der Tonhöhenverteilung zu erkennen: Die breite Verteilung bei F. W. geht in den Formantbereich über (Zeichen für Vitalität), während bei den anderen beiden Sprechern Kumulationen bei 125 bzw. 250 Hz zu

beobachten sind, was auf eine mehr monotone Sprechweise schließen läßt. Die Proben von anderen Abschnitten des gelesenen Textes ergeben Deckungsgleichheit der Kurven.

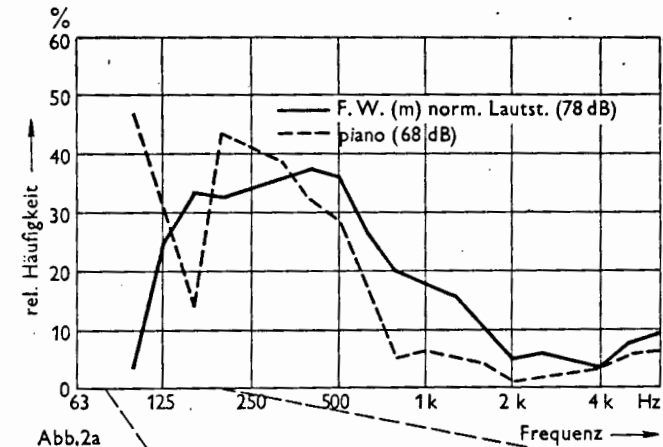


Abb. 2a

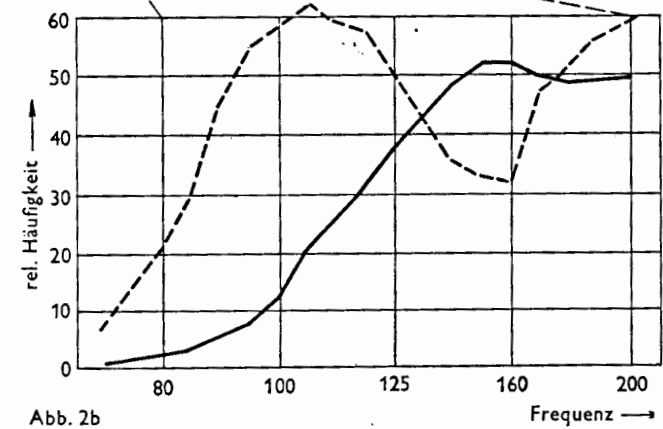


Abb. 2b

Neben diesem prosodischen Merkmal wird ein Hinweis auf die Lautstärke bzw. Dynamik durch den Versuch gegeben, normal und dann denselben Text lauter zu lesen (Abb. 2a). Der Flächenunterschied zwischen den beiden Kurven ergibt ein Maß für Lautstärke bei gleichzeitig veränderter Intonation.

Es sind ferner noch folgende Sprechweisen untersucht worden: zu schnell, zu langsam, zu hoch, ferner Störungseinflüsse auf den Sprecher wie Geräusch und Lee-Effekt, sprechen mit und ohne Zähne, Gesangsstimmen, Stottern und sonstige Stimm- und Sprachfehler.

Eine gewisse Grundtendenz ist darin zu erblicken, daß jede zusätzliche Beanspruchung, sei es durch Erschwerung des Sprechaktes selbst oder durch Ablenkung der Aufmerksamkeit (Echo oder Lärm), die Artikulationstiefe bzw. den Artikulationsfaktor reduziert. Die Frequenzhäufigkeitsverteilung tendiert allerdings nicht

in dem erwarteten Maß zur Kumulation in den Formantbereichen des neutralen Vokals. Die Terzfilteranalyse ist dafür außerdem zu grob. Ferner führt das Nach-

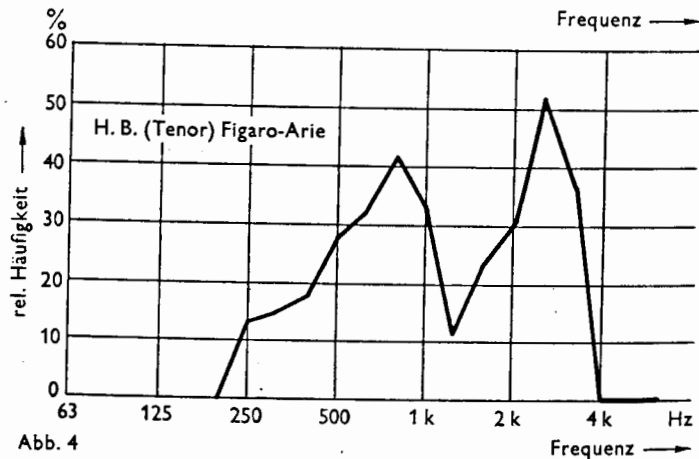
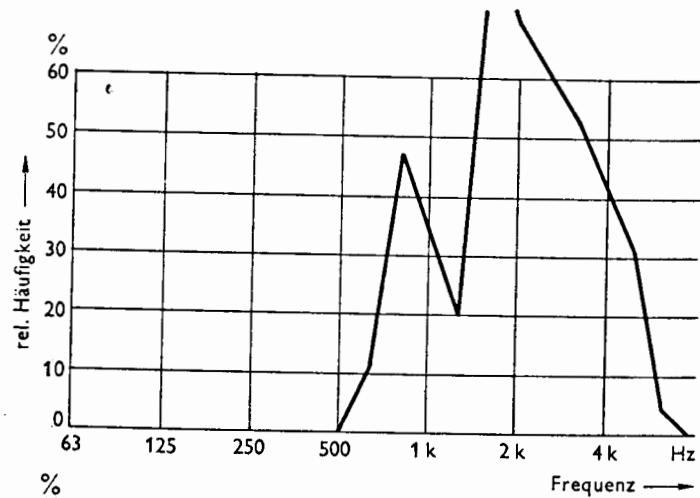
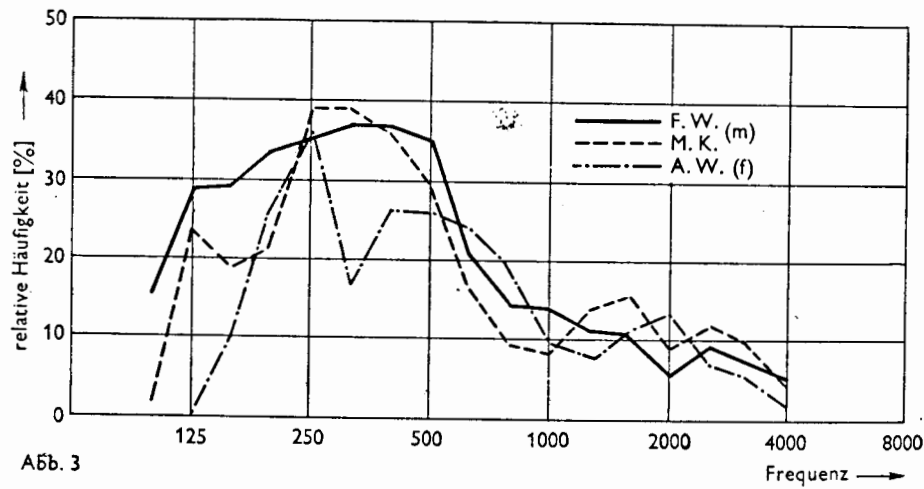


Abb. 4

lassen der Spannung in der Artikulationsmuskulatur zur Amplitudenverringerng des Formantbereichs um 2500—3500 Hz (Vokal i). Eine erhöhte Spannung in diesem Bereich ist bei ausgebildeten Sprechern stets vorhanden.

Weiterhin bewirken die zusätzlichen Beanspruchungen beim Sprechen durch Störungseinflüsse eine Verringerung des Intonationsumfangs im Sinne zunehmender Monotonisierung. Die Individualunterscheidung beim normalen Sprechen drückt sich in den Tonhöhen-Diagrammen durch den Grad der Monotonie aus, außerdem in der Amplitude des 3000-Hz-Formantbereichs. Bei lebhaften Sprechern wird der Einschnitt zwischen Tonhöhen- und Formantbereich überbrückt wegen der breiteren Skala der Melodiehöhe.

Aus dem Diagramm „Tenorstimme — Hahnruf“ (Abb. 4) geht eine weitgehende Übereinstimmung der Häufigkeitsmaxima hervor. Daraus läßt sich die Erkenntnis ableiten, daß in der spektralen Häufigkeitsverteilung wesentlich die Charakteristik des Stimmerzeugers vorhanden ist. Man kann dann in dem gezeigten Beispiel der ausgeprägten Formantspitzen, die im Verhältnis 2 : 1 und 3 : 1 vorkommen, auf die Form der Generatorschwingung schließen, im letzteren Fall also auf eine Rechteckschwingung, was im Prinzip auf eine andersartige Lösung gegenüber dem Verfahren des „Inverse Filtering“ hinausläuft.

LITERATUR

F. Winkel u. M. Krause, Ermittlungen von Spracheigenschaften aus statistischen Eigenschaften von Amplitude und Tonhöhe, *Int. Akustik-Kongreß Lüttich (A 41) 1965*.

SPRECHSTÖRUNGEN BEI PARKINSONPATIENTEN

HENNING WODE*

An der Universität Freiburg i. Br. untersuchen wir z. Z. in Zusammenarbeit mit der Neurochirurgischen Universitätsklinik die Sprechstörungen bei deutschsprachigen Parkinsonpatienten. Aus dem Gesamtkomplex der Sprechstörungen möchte ich in rein linguistischer Sicht zwei Fragen herausgreifen: 1. die Frage nach dem System — welche Schicht der Sprachstruktur wird angegriffen? und 2. ob sich eine allgemeine Verfallsrichtung erkennen läßt oder ob die Störungen ohne derartige Gesetzmäßigkeiten auftreten.

Nach der Untersuchung von rd. 60 Patienten scheint folgende Hypothese gerechtfertigt: In Mitleidenschaft gezogen ist allein die phonetische Schicht der Sprachstruktur, genauer: die Aussprache. Das gesamte System der Sprache ist abgesehen von der lautlichen Realisierung der linguistischen Einheiten völlig ungestört erhalten.

Für die Annahme, das System selbst sei ungestört, sprechen u. a. folgende Beobachtungen: 1. Die Patienten verstehen alles, selbst Erläuterungen zu Dingen, die ihnen völlig neu sein müssen. So führen sie z. B. nach mündlicher Anweisung krankengymnastische Übungen aus. Ähnlich wurden die Erklärungen zu den linguistischen Tests einwandfrei verstanden.

2. Selbst solche Patienten, die in freier Unterhaltung weitgehend unverständlich reden, sprechen einzelne Wörter befriedigend nach, und zwar in der Regel nicht in der Mundart, wie sie ihnen vorgesprochen wurden, sondern in ihrer eigenen. Das kann nur heißen, das vorgesprochene Wort wird zunächst dekodiert und dann gemäß dem eigenen Dialekt kodiert und ausgesprochen. Dazu muß das System, soweit es für den jeweiligen Fall relevant ist, intakt sein.

3. Die Zuordnung von suprasegmentalen Elementen (Betonung, Tonhöhe, Pausen) zu syntaktischen Strukturen und ihren Manifestationen ist ungestört. Wir haben das überprüft an Hand von rd. 40 Morphemfolgen, bei denen die jeweilige Morphemkette mehrere syntaktische Interpretationen zuläßt, die in ihrer Normalintonation¹ auch hörbar unterschieden werden. z. B.

* Englisches Seminar der Universität Freiburg i. Br.

¹ Vgl. Verf., Englische Satzintonation, *Phonetica* 15 (1966), 129—128.

(1a) Fürchtegott //²

(1b) fürchte Gótt //

(1a) ist ein Vorname, (1b) die syntaktische Gruppe Verb + Objekt.

(2a) Herr / im Himmel steh mir bei //

(2b) Herr im Himmel / steh mir bei //

In (2a) ist die Morphemfolge *im Himmel* Adverb, in (2b) gehört sie zur Anrede und erweitert den Kern *Herr*.

4. Selbst bei segmental völlig unverständlichen Äußerungen kann die Intonation erhalten sein, so daß man z. B. glaubt, eine Aufzählung, eine Frage usf. vor sich zu haben.

5. Mir sind nie Fehlleistungen bei der Flexion, der Kongruenz, im Satzbau, bei der Wortwahl usf. aufgefallen, die über die bei gesunden Sprechern auch vorkommenden hinausgehen. Auch verständliche Fetzen in sonst unverständlichen Äußerungen wiesen keine derartigen Mißbildungen auf.

Die Sprechstörungen bei Parkinsonpatienten sind gradueller Art. Betrachtet man die Auswirkungen der Veränderungen auf die Realisierung der distinktiven segmentalen Elemente, so lassen sich in unserem Material zwei Verfallstypen unterscheiden:

Typ 1: Abbau der Stimmhaftigkeitskorrelation

Typ 2: Abbau der Strikturen oberhalb der Glottis

Diese beiden Typen sind in der Weise unabhängig voneinander, daß das Vorhandensein des einen Typs notwendigerweise das Fehlen des anderen bei demselben Sprecher impliziert.

Bei Typ 1 wird die Stimmhaftigkeit mehr und mehr verringert, bis die Patienten schließlich über weite Strecken völlig stimmlos sprechen. Dabei tritt bisweilen Flüsterreibung auf. Unter unseren Sprechern erfolgte der Abbau der Stimmhaftigkeit nach folgender Regel: Solange der Silbenkern stimmhaft war, konnten fakultativ auch andere Konstituenten der Silbe stimmhaft sein. War der Kern stimmlos, blieb die ganze Silbe stimmlos. Eine Verständigung ist in diesen Fällen u. a. infolge der Redundanz gut möglich.

Beim Typ 2 werden die Strikturen oberhalb der Glottis in Richtung auf eine neutrale Artikulationsstellung abgebaut mit leichter, schwer zu lokalisierender Reibung. Die Lokalisierungskorrelationen sind (weitgehend) aufgehoben. Die Lippen bleiben ungerundet. Die Zungenbewegung wird auf eine zentrale Stellung zwischen offen und halboffen hin mehr und mehr eingeschränkt.

Im einzelnen werden zu erwartende Verschlußstrikturen zunehmend schwächer geschlossen und erscheinen schließlich als spirantische Strikturen. Betroffen werden die oralen Verschlüsse, sowie [r l] etc.

Spirantische Strikturen werden mehr und mehr geöffnet. Die Reibung nimmt gleichzeitig ab. Spirantische Strikturen, die für zu erwartende Verschlüsse erscheinen, haben geringe oder fast gar keine Reibung.

Die Vokale werden in Richtung auf eine zentrale Zungenstellung zwischen offen und halboffen abgebaut.

Als Endpunkt des Verfalls nach Typ 2 ist uns mehrfach eine unverständliche Sprechweise begegnet, in der die Merkmale geschlossen ≠ geöffnet, Reibung ≠ Nichtreibung, oral ≠ nasal, Konsonant ≠ Vokal sowie die Lokalisierungskorrelationen nicht mehr identifizierbar waren.

N.B. Zum Stand der Untersuchungen bis 1969 vgl. Verf., *Linguistische Untersuchungen zum Parkinsonismus*, Basel (Karger) 1969.

DISCUSSION

Holm:

Sie sprachen über den Sprachabbau. Ich fürchte überhört zu haben, ob die Untersuchungen vor oder nach einer Operation durchgeführt wurden, da die Verhältnisse prä- und postoperativ doch verschieden sind, und uns so weitere Aufschlüsse über den zentralen Mechanismus geben können.

Wode:

Meine Ausführungen betrafen nur präoperative Verhältnisse. Wir untersuchen gleichzeitig auch die postoperativen Störungen. Bislang deutet alles darauf hin, daß die Operation keinen Einfluß auf die Verfallstypen hat, d. h. auch postoperativ ersetzen die Kranken spirantische Strikturen nicht etwa doch Verschlüsse. Ändern kann sich hingegen der Schweregrad der Störungen. Die Sprechweise kann im Rahmen des betr. Verfalltyps besser oder schlechter werden oder unverändert bleiben.

² betont, unbetonte Silben werden nicht markiert. / Innenpause, // Endpause. Vgl. Verf., *Pause und Pausenstelle im Deutschen*, Acta linguistica Hafniensia (im Druck).

SOME PHONETIC PHENOMENA IN NEW GUINEA LANGUAGES

S. A. WURM

Some of the very numerous Papuan (or non-Austronesian) languages of the New Guinea area are characterized by the occurrence in them, of speech sounds which are uncommon in languages in other parts of the world.

One of these sounds is a laterally released velar stop or more correctly a heterorganic affricate, which is typical of languages of the so-called Central and East-Central Families situated in the Highlands Districts of Australian New Guinea and also of languages located in the Huon Peninsula. While it has phonemic status in all the languages in which it occurs, its principal allophonic forms in different languages are phonetically of two distinct types. One of them is a usually voiced, midvelar stop with a relatively long fricative lateral release with the tongue-tip usually articulating against the alveola. The other type is a post-velar, always voiced stop with a relatively short, heavy-friction lateral release with the tongue-tip pointing forward without touching anywhere. The former is more common, the latter being restricted to the Yagaria language of the East-Central Family in which it occurs with great statistical frequency. Characteristically, a phonologically conditioned allophone of the former is a voiceless lateral fricative, whereas the latter has no such allophone, and little allophonic variation in general.

Other unusual sounds, also occurring in Yagaria, are pre-glottalized voiced bilabial and alveolar stops. They occur only word-medially and are in complementary distribution with non-preglottalized voiced bilabial and alveolar stops, e.g. [^ldε^ʔdæε] 'they ate'. For phonological reasons, these preglottalized stops are interpreted as complex phonetic units rather than as CC sequences: while preglottalization may occur word-medially with all consonants except the voiceless stops, b and d are never found without preglottalization in medial position; therefore, other preglottalized medial consonants are phonemically interpreted as CC, whereas [^ʔb] and [^ʔd] are regarded as allophones of [b] and [d] with the allophones [b] and [d] occurring initially.

An unusual sound found in parts of the Southern Highlands is a nasalized trill [ɾ̃] which appears in contrast with a non-nasalized trill [ɾ], e.g. Foe [*gakaro*] = 'short' as opposed to [*yeṛahae*] = 'stand up'. This sound is similar to one found in the Ponérihouen Language in New Caledonia, e.g. [*maṛu*] = 'bird'. In languages of the same area, and of adjacent parts of the Western District, nasal and non-nasal flaps are

encountered in at least one of the languages. Their exact phonemic status is still under study. These sounds have been found both as ordinary and as retroflexed flaps.

In the Kilmeri Language, in the north-western part of the Sepik District, a bilabial voiced trill occurs medially, apparently as an allophone of (b).

The existence of tonal languages in New Guinea has been known for many years, but until not much more than a decade ago, they had been thought to be exceptional. With the rapid expansion and development of New Guinea Linguistics in recent years, an increasing number of languages believed to be tonal has been reported. While in some instances, like with Yabem on the coast of the Huon Peninsula, and Telefol in the centre of the island, this assumption appears to be correct, there have been instances in which different linguists working in two closely related languages (like for instance Gahuku and Benabena in the Eastern Highlands) analyzed one of them as tonal and the other one as non-tonal with phonemic stress conditioning syllable pitch. While this is certainly a possibility, it suggests that there may be room for different interpretations of similar or identical supra-segmental phenomena in related languages. Recent studies have shown that in a large number of New Guinea languages, especially in the Highlands languages of Australian New Guinea, supra-segmental phenomena are present which on first impression appear to be syllable tones, but which upon careful analysis can be explained differently. So for instance, the postulation of three emic stress levels in Yagaria permits the interpretation of tonal phenomena as pitch changes conditioned by stress, e.g. ($1h\acute{a}^3t\bar{i}$) = 'five', [$^3h\acute{a}^2li$] = 'arrow'. Such cases seem to be common in the Highlands, and recent work by S. Wurm in languages of the Trans-Fly area in the Western District has shown that a similar situation appears to exist with some languages there. So-called 'ornamental' tones have been mentioned for other parts of New Guinea, and it seems that languages with complex interacting supra-segmental systems are quite usual in New Guinea.

DISCUSSION

Carnochan:

From your description and from listening to your recorded examples, I think that articulations very similar to your pre-glottalized plosives occur in a number of Nigerian languages such as Fula, Hausa and Bachama e.g. in Hausa "Sun babbazgi itace" they uprooted all the trees "Sun daddala mashi sata", the pre-glottalization and voicelessness is particularly noticeable in the long stops corresponding to the double letters.

REGELN FÜR DIE FREMDWÖRTERBETONUNG IM DEUTSCHEN

W. U. WURZEL

Es ist bekannt, daß der Wortakzent im Deutschen im allgemeinen leicht zu beschreiben ist und dem Linguisten wenig Schwierigkeiten bereitet. „Im allgemeinen“ bedeutet hier allerdings: bei den [+ Native]-Wörtern.

Der Wortakzent der deutschen Fremdwörter oder [—Native]-Wörter macht dagegen auf den ersten Blick einen recht willkürlichen und unregelmäßigen Eindruck, vgl. z. B. die Betonung von 'Syndikus (1. Silbe), Pi'rat (2. Silbe), Dema'goge (3. Silbe), Ökonomie (4. Silbe) usw. oder das Nebeneinander von Mu'sik, 'Musiker, musi'kalisch und 'Logik, 'Logiker, 'logisch. Daneben gibt es noch Fälle wie Te'nor—'Tenor und Par'tie — 'Party. Der Eindruck der Regellosigkeit wird durch Nachschlagen in Grammatiken nur verstärkt. So heißt es beispielsweise bei Jung: „Wir wiederholen aber nochmals: Für die Betonung der Fremdwörter lassen sich schwer eindeutige, feste Regeln geben. Für den Einzelfall muß auf die Aussprache- und Fremdwörterbücher verwiesen werden“¹ und: „Maßgeblich für die Wortbetonung sind Ort und Zeit der Übernahme und die Häufigkeit des Gebrauchs.“²

Was uns hier geboten wird, sind alles extralinguistische Kriterien soziologischer Art, die nicht in eine formalisierte Grammatik, ja eine Grammatik überhaupt zu integrieren sind. Heißt das nun, daß es keinerlei innerlinguistische, phonologische Kriterien für die Festlegung des Akzents bei [-Native]-Wörtern gibt, wenigstens keine einheitlichen? Werden diese Wörter einfach so betont wie in ihren Herkunftssprachen, so daß im Grunde eine — über wieviel Zwischenstufen auch immer vermittelte — Kenntnis der Akzentverhältnisse der Herkunftssprachen notwendig ist? Dafür scheinen Beispiele wie Par'tie — 'Party zu sprechen.

Gegen diese Annahme sprechen die folgenden Fakten:

(i) Ein konstruiertes Wort, das aufgrund seiner segmentalen Struktur als [-Nativ] aufgefaßt wird, wird auf An-hieb einheitlich („richtig“) betont, so z. B. nārīma als 'nārīma, narīma als na'rīma und nārīmān als nārī'mān.

(ii) Oft stimmen die deutsche [-Nativ]-Betonung und die Betonung in der Herkunftssprache nicht überein, vgl. z. B. russ. Koma'rov — dtsh. Ko'marov, tsch. 'Kač-

¹ Jung, Grammatik der deutschen Sprache, Leipzig 1966, S. 481. Ähnlich auch in: Wörterbuch der deutschen Aussprache, Leipzig 1964, S. 81.

² W. Jung, op. cit. S. 480.

mārek — dtsh. *Katsch'mārek*, tsch. 'Katinka — dtsh. *Ka'tinka*, isld. 'berserkur — dtsh. *Ber'serker* und isld. 'geysir — dtsh. *Gey'sir* (neben 'Geysir). Es kommt also nicht auf die Betonung in der Herkunftssprache an, auch wenn die deutsche Betonung oft nicht mit ihr übereinstimmt. Das Entscheidende ist vielmehr die Silbenstruktur. Es gibt drei für unser Problem relevante Arten von Silben: „überschwere“ Endsilben, die stets betont werden, „schwere“ Silben, die betont werden, wenn es keine „überschweren“ Silben gibt, und „leichte“ Silben, die wiederum betont werden, wenn es weder „überschwere“ noch „schwere“ gibt.³ Wenn wir das berücksichtigen, können wir in folgender Weise die richtigen Wortakzente der Fremdwörter im Deutschen erzeugen:

(i). Aufgrund des Merkmals [α gespannt] bei den Vokalen und der dem entsprechenden Vokal folgenden Konsonantencluster wird die Anschlußart prädiert und damit die Silbengrenzmarkierung eingeführt. Das geschieht durch einige nicht sehr komplizierte Regeln. Regeln dieser Art sind aus unabhängigen phonetischen Gründen ohnehin notwendig.

(ii) In einer weiteren Regelgruppe wird die Betonbarkeit bzw. Nichtbetonbarkeit festgelegt. Unbetonbar sind die Endungen -*ön*, -*üm* und -*üs* und alle [+ Nativ]-Suffixe (außer -*ei*), die sich natürlich auch mit [-Nativ]-Elementen verbinden können. Diese Regeln sind sehr einfach. Da sie eine geordnete Menge darstellen, kann ihnen eine Regel folgen, die alle übrigen Vokale als betonbar charakterisiert.

(iii) Jetzt wird die eigentliche Betonungsregel für [Nativ]-Wörter angewandt:

$$\left[\begin{array}{l} + \text{vokal} \\ - \text{kons} \\ + \text{B} \end{array} \right] \rightarrow [1 \text{ bet}] \mid \left\{ \begin{array}{l} \text{a) } \left\{ \begin{array}{l} [+ \text{lang}] \\ [- \text{lang}] \end{array} \right\} \left\{ \begin{array}{l} [- \text{vokal}]_0 \\ [- \text{vokal}]_2 \end{array} \right\} \# \\ \text{b) } \left\{ \begin{array}{l} [+ \text{lang}] \\ [\alpha \text{lang}] \end{array} \right\} \left\{ \begin{array}{l} [- \text{vokal}]_1 \\ [- \text{vokal}]_1 \end{array} \right\} \mid X \# \\ \text{c) } [\alpha \text{lang}] \left\{ \begin{array}{l} [- \text{vokal}]_0 \\ [- \text{vokal}]_0 \end{array} \right\} \mid X \mid Y \mid Z \# \\ \text{d) } [\alpha \text{lang}] \left\{ \begin{array}{l} [- \text{vokal}]_0 \\ [- \text{vokal}]_0 \end{array} \right\} \mid X \mid Y \# \\ \text{e) } [- \text{lang}] \left\{ \begin{array}{l} [- \text{vokal}]_1 \\ [- \text{vokal}]_1 \end{array} \right\} \# \end{array} \right.$$

'#' kennzeichnet die Wortgrenze, '|' die Silbengrenze. X, Y, Z enthalten keine Silben — bzw. Wortgrenzen, [+ B] charakterisiert die Betonbarkeit, [1 bet] den Wortakzent.

Regel (a) betont die „überschweren“ Endsilben (Silbenstruktur: $-\bar{V}$, $-\bar{V}K$, $-\bar{V}K_2^2$), z. B. *Ökono'mie*, *Par'tie*, *Mu'sik*, *Ga'rant*, *So'list*. Regel (b) erfaßt die Betonung der

³ Es kommt hier nicht auf die vielleicht nicht besonders glücklichen Bezeichnungen, sondern auf die verschiedenen relevanten Silbenstrukturen an, wie sie die Betonungsregel erfaßt.

⁴ Die „überschweren“ Endsilben unterscheiden sich nur in einem Punkt von den „schweren“, nämlich daß sie Cluster der Struktur $-\bar{V}K$ nicht einschließen. Dagegen sind Endsilben mit dem Ausgang $-\bar{V}$ als „überschwer“ zu fassen (wegen *Ökonomie*, *Par'tie* usw.). Diese könnten allerdings auch mit $-\bar{V}V$ (vgl. Orthographie) repräsentiert werden.

„schweren“ vorletzten Silben (Struktur: $-\bar{V}$, $-\bar{V}K_1$) z. B. *öko'nomisch*, *musi'kalisch*, *Dema'goge*, 'Logik, 'Tenor.

Regel (c) weist Vier- und Mehrsilbern, die weder durch (a) noch durch (b) erfaßt werden, die Betonung auf der viertletzten Silbe zu, und zwar unabhängig von ihrer Silbenstruktur: 'Omnibusse, 'Syndikusse, musi'kalischen, Regel (d) tut dasselbe mit den noch nicht erfaßten Dreisilbern, die auf der 1. Silbe akzentuiert werden; vgl. z. B. 'Omnibus, 'Syndikus, 'Korridor, 'Januar.

Regel (e) schließlich ist notwendig für Fälle wie *Ma'jor*, *Te'nor*, *Tre'sor*, d. h. für noch nicht erfaßte Wörter mit „schweren“ Endsilben (Struktur: $-\bar{V}K$). Wie schon bei *Korridor* und *Januar* sind hier die Endsilbenvokale nicht phonemisch sondern nur phonetisch sekundär lang. Die Dehnung erfolgt unter dem Einfluß des -r (vgl. *dēr* — *des* usw.). Aus dem gleichen Grunde heißt es auch nicht **Dok'tör* (nach (a)), sondern 'Doktör (nach (b)). „Echt“ lang ist das o nur im Plural, deshalb *Dok'tören*.⁵

(iv) Abschließend folgen noch 2 Regeln, die die Dehnung kurzer betonter Vokale und die Kürzung langer unbetonter Vokale in offener Silbe bewirken. Diese Regeln gestatten die Ansetzung einheitlicher Basisformen für z. B. *Müse* und *Musik*, *Lōgik* und *Logistik* sowie *Nākleus* und *nuklear*: /muz/, /log/ und /nukle/.

Sie sehen folgendermaßen aus:

Dehnungsregel (2)

$$[- \text{lang}] \rightarrow [+ \text{lang}] \quad / \quad \left[\begin{array}{l} + \text{vokal} \\ - \text{kons} \\ 1 \text{ bet} \end{array} \right]$$

Kürzungsregel

$$[+ \text{lang}] \rightarrow [- \text{lang}] \quad / \quad \left[\begin{array}{l} + \text{vokal} \\ - \text{kons} \\ 0 \text{ bet} \end{array} \right]$$

Auf Beispiele für ihre Wirkung müssen wir hier leider wieder verzichten.

⁵ Wörter bzw. Morpheme dieser Art werden im Lexikon mit Kurzvokalen repräsentiert, [*doktor*], [*cäzar*], [*elektron*], [*dämon*]. Vor der Betonungsregel wird eine Dehnungsregel angewandt, die kurze Vokale vor Silbengrenze und folgendem ungehemmten Konsonant [*l, n, m, r*] längt:

Dehnungsregel (1)

$$\left[\begin{array}{l} + \text{vokal} \\ - \text{kons} \\ - \text{lang} \end{array} \right] \rightarrow [+ \text{lang}] \quad / \quad \left[\begin{array}{l} + \text{kons} \\ - \text{geh} \end{array} \right]$$

Diese Regel beeinflußt *Doktor*, *Cäsar*, *Elektron* und *Dämon* nicht, dehnt aber z. B. die Vokale der Pluralformen: *Doktören*, *Cäsären*, *Elektrönen*, *Dämönen*. Es folgt dann die Betonungsregel. Erst viel später im Ableitungsmechanismus operiert dann eine weitere Dehnungsregel (3), die eine Art „sekundäre“ (oder „tertiäre“) Vokallänge vor ungehemmten Konsonanten bewirkt. Sie scheint jedoch Nichtnasale und Nasale nicht gleichmäßig zu erfassen, vgl. [*doktör*] vs. [*dämon*], dag. [*dēr*] vs. [*dēn*]. Diese Probleme, die mit dem Wortakzent eng zusammenhängen, bedürfen noch einer eingehenden Untersuchung.

Damit haben wir die richtigen Wortakzente für nahezu alle deutschen [— Nativ]-Wörter erzeugt. Nur einige wenige Wörter dieser Art entziehen sich den formulierten Regularitäten (z. B. *Leutnant*). Hier wird offensichtlich der [— Nativ] - Akzent. Interessanterweise erfaßt die hier diskutierte Betonungsregel nicht nur Wörter fremder Herkunft, sondern auch einige Wörter deutscher Herkunft, die aufgrund ihrer Struktur als [—Nativ] aufgefaßt werden: *Fo'relle*, *Ho'lunder*, *Hor'nisse*, *Wach'holder*, *Herme'lin*. Zusammenfassend läßt sich also konstatieren, daß die Betonung der Fremdwörter im Deutschen nicht durch irgendwelche schwer zu fassende fremdsprachlichen oder soziologischen Fakten determiniert werden muß und kann. Sie funktioniert nach innerphonologischen deutschen Regularitäten, bei deren Entstehung fremdsprachliche Fakten (z. B. französische und lateinische) wohl eine Rolle gespielt haben, aber von denen sie heute unabhängig ist.

DISCUSSION

Martens:

Der Ansatz scheint sehr fruchtbar. Bei einigen Beispielen sind die als „nicht-vorkommend“ eingeordneten Akzentdistribution jedoch sehr verbreitet, z. B. „Berserker“ und „Katschmarek“ mit Akzent auf der 1. Silbe. — Außerdem sollte nicht unbeachtet bleiben die Tendenz zur Anpassung an das Prinzip der gleichartigen Akzentverteilung bei Singular und Plural. Von daher löst sich das Problem „Motor-Motoren“ — mit Akzent auf der 2. Silbe — nämlich anders auf.

KYBERNETISCHE SPRECHMODELLE

LUDWIK ZABROCKI

In unserem Referat werden wir uns bemühen, nur ein Teilgefüge des sprachlichen Kommunikationsgefüges darzustellen. Das dargestellte Modell werden wir als ein zu approximierendes Modell betrachten. Als beliebiges volles kybernetisches Grundgefüge betrachten wir ein abstraktes Informationsmodell, daß sich unter anderem dadurch charakterisiert, daß sämtliche Kanäle 3 Kreise enthalten und, zwar 1. ein Grundinformationskreis, 2. einen Kontrollkreis, 3. einen Steuerungskreis. Diese Einheiten werden von uns als konstitutive Faktoren eines kybernetischen Gefüges betrachtet, das als Basis für den Umlauf von Informationsströmen dient. Neben Kanälen ist jedem kybernetischen Gefüge ein System von Transformationsknoten verschiedener Art eigen. Es sind transponierende und datenverarbeitende Transformationsknoten. Die Dreischaltungskreisstruktur der Kanäle wird von Zentraltransformationsknoten determiniert.

Den Grundinformationskreis kann man auch in bezug auf den Empfänger als Grundsteuerungskreis bezeichnen. Dagegen sind der Kontrollkreis sowie der Steuerungskreis im Verhältnis zum Grundinformationskreis (oder dem Grundsteuerungskreis) Metakreise. Sie sind mit der Auslösung, dem Verlauf und der Auswirkung der Grundinformation verbunden. Eng miteinander gekoppelt bilden sie Kreisrelationen. Die Grundinformation wird von ihnen kontrolliert und gesteuert. Das sprachliche Informationsgefüge charakterisiert sich noch dadurch, daß der Grundinformationskreis Träger einer systemaufbauender Funktion ist. Das bedeutet, daß im Zentraltransformationsknoten, daß heißt letzten Endes im Gehirn eine Art von Sprachstrukturmatrizen sowie Sprachspeicher aufgebaut werden.

Das Sprechgefüge als Teil eines Sprachkommunikationsgefüges besteht aus einem Sender und einem Empfänger, weiter aus Kanälen, die die lautliche Substanz in adäquat transformierter Form sowie in adäquat retransformierender Form übertragen. Graphisch könnte man das folgendermaßen darstellen:

Das Sprechgefüge zerfällt in drei Subgefüge. Das erste Subgefüge besteht aus einem Kanal, der das Sprachzentrum im Gehirn mit den Artikulationswerkzeugen verbindet, weiter führt dieser Kanal vom Munde des Senders zu Gehörsrezeptoren des Empfängers und von dort zum Sprachzentrum des Empfängers. In dem Kanal α können beliebige Transformationsknoten eingebaut werden. Es können aber nur transponierende Transformationsknoten sein. Man kann auch die Grundinformation in

diesem Kanal speichern (Grammophon, Tonbandaufnahmen usw.). Auch in diesem Falle haben wir es nur mit transponierenden Transformationsknoten zu tun, die aber zugleich Speicherknoten sind.

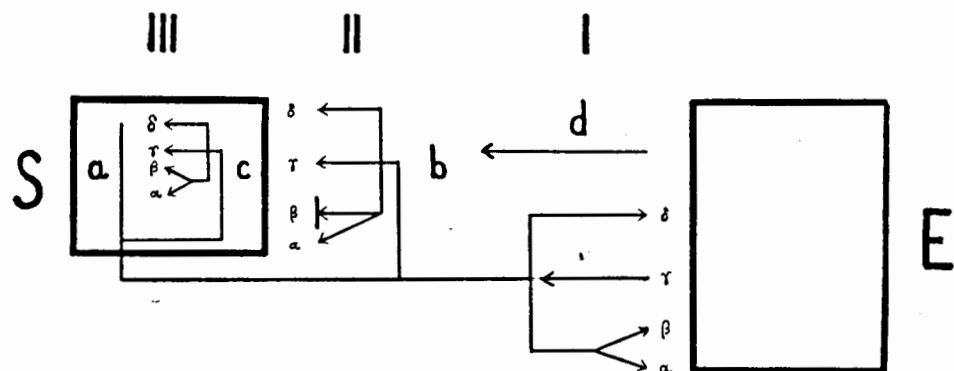


Abb. 1. Erklärung: S = Sender, E = Empfänger, I = erstes Subgefüge, II = zweites Subgefüge, III = drittes Subgefüge, α = Grundinformationskreis, β = Grundinformationskreis als Aufbaukreis, γ = Kontrollkreis, δ = Steuerungskreis, β = blockierter Aufbaukreis; a, a_1, a_2 — Hauptkanal, b — Abzweigungskanal des II Subgefüges, c — Abzweigungskanal des III Subgefüges, d — optischer Kontrollkanal im I Subgefüge.

Der Informationskanal a des ersten Subgefüges enthält, wie jedes grundkybernetisches Gefüge, drei Schaltungskreise, nämlich den Grundinformationskreis, den Kontrollkreis und den Steuerungskreis. Der Grundinformationskreis ist zugleich Träger einer systemaufbauenden Funktion. Diese Funktion wird vom Empfänger in dem Sinne determiniert, daß der Empfänger bei der Verarbeitung der Grundinformation gleichzeitig das Produktionssystem, auf Grund dessen die Grundinformation beim Sender ausgelöst worden ist, adäquat in seinem Sprachzentrum aufbauen kann.

Der Aufbaukreis wird nach Aufbau von sprachlichen Strukturmatrizen teilweise abgebaut. Praktisch geschieht das vom 15. Lebensjahr an. Das Absterben des Aufbaukreises setzt zu dem Zeitpunkt an, wo dem entsprechenden Strukturmatrizen-aufbau Informationen gleich Null zugesteuert werden. Das geschieht zuerst im Bereiche der Phonetik, dann der Morphologie sowie endlich der Syntax. Der Aufbau von Wort- und Syntagmaspeichermatrizen ist das ganze Leben hindurch tätig. Auf Grund dieser Tatsachen wird der Altersparameter beim Fremdsprachenunterricht aufgebaut und entsprechend ausgewertet.

Die Grundinformation sowie die Metasteuerungsinformation fließen vom Sender zum Empfänger. Die Kontrollinformation dagegen muß in umgekehrter Richtung fließen. Falls sie optischer Natur ist, fließt sie im Kanal b . Meistens ist sie jedoch phonischer Natur. In diesem Falle fließt sie im Kanal a in der Weise, daß in unserem Modell der Empfänger zum Sender wird.

Die Kontrolle der adäquaten Perzeption der sprachlichen Mitteilung durch den

Empfänger wird auf Grund der Antwort des Empfängers, also auf dem Wege über den Lautkanal durchgeführt. Die Kontrolle kann auch auf Grund des allgemeinen Verhaltens des Empfängers, also durch den optischen Kanal, durchgeführt werden. Der Kontrollkreis operiert somit mit phonischen und optischen Signalen. Die Kontrolle kann zeitlich verschieden sein. Es können Bruchteile von Sekunden, Minuten, Stunden, ja, sogar Jahre vergehen, ehe eine Kontrolle zustande kommt. Im äußersten Falle wird die Kontrollfunktion des Informationskanals nicht realisiert. Wir können hier von einer Neutralisation der Kontrollfunktion sprechen.

Neben dem soeben beschriebenen Teilkommunikationsgefüge, das wir als erstes Subgefüge bezeichneten, gibt es, wie schon eingangs betont wurde, noch ein zweites Subgefüge. Dieses Subgefüge besitzt mit dem ersten Subgefüge teilweise gemeinsame Kanäle, teilweise jedoch auch einen eigenen Kanal. Vom Sprachzentrum bis zu den Sprachwerkzeugen teilt es mit dem ersten Subgefüge dieselben Kanäle, dagegen im weiteren Verlauf wird der gemeinsame Kanal gespalten. Der abgezweigte Kanal führt vom Munde des Senders zu den Gehörsrezeptoren des Senders. Es ist somit ein internes Subgefüge. Wir bezeichnen es als zweites Subgefüge, die einzelnen Kreise als zweite Schaltungskreise. Die Hauptfunktion fällt hier dem Kontrollkreis zu. Überhaupt kann man dieses Subgefüge als ein ausgesprochenes Kontrollgefüge bezeichnen. Es kontrolliert nämlich die konkrete Produktion der phonischen Signale. Die Kontrolle wird auf allen Ebenen des konkreten Sprechaktes durchgeführt. Somit werden u. a. kontrolliert: 1. die phonische Substanz, 2. die grammatikalischen Strukturen, 3. die Wortwahl. Der gesamte Sprechakt wird hier auf die Adäquatheit der Reproduktion von paradigmatischen Einheiten sowie auf die Produktion von syntagmatischen sprachlichen Einheiten kontrolliert. Dieser Kreis ist von größter Bedeutung für die Erlernung einer Sprache.

Im zweiten Subgefüge wird der Aufbaukreis blockiert. Dieser Kreis funktioniert im zweiten Subgefüge nicht. Somit wird bei der Spracherlernung die eigene fehlerhafte Aussprache nicht zum Aufbau von Strukturmatrizen zugelassen. Unsere Sprachmatrizen sind somit nicht auf Grund unserer eigenen Aussprache aufgebaut, sondern auf Grund der Sprache der Umgebung. Diese Tatsache dürfte von entsprechender Wichtigkeit für die experimentelle Phonetik sowie für den Fremdsprachenunterricht sein.

Im Kanal des zweiten Subgefüges können keine zusätzlichen Transformationsknoten eingebaut werden.

Die korrekte Funktion dieses Kontrollkreises bildet wohl die Grundlage für die blockadefreie Steuerung des Stromes der Sprechimpulse im Sprachzentrum des Gehirns, das heißt, daß sie das Stottern verhindert.

Es gibt noch ein drittes Subgefüge. Dieses basiert auf dem ersten Subgefüge auf der Strecke vom Sprachzentrum bis zu den Sprachwerkzeugen. Von dort führen entsprechende Nervenkanäle zurück zum Sprachzentrum. Es ist wieder ein internes Subgefüge. Dieses Subgefüge ist am wenigsten erforscht. Mit diesem Gefüge hat sich unlängst Professor Winckel befaßt (*Phonetica*, Vol. 9, No. 2, 1963, S. 108—126).

Man kann annehmen, daß hier mit Ausnahme des Aufbaukreises alle Schaltungskreise funktionieren. In pathologischen Fällen wird wohl auch der Aufbaukreis tätig sein.

DISCUSSION

Martens:

Die Darlegung liefert erneut den Beweis, daß beim Sprachunterricht eine Periode des Hörenlernens der Periode des Sprachenlernens vorausgehen sollte, also zuerst Hörübungen — unserer Meinung nach allerdings von vornherein auf das phonologische System der Zielsprache ausgerichtet! — dann Ausspracheübungen (im weitesten Sinne).

FORM-FUNKTIONSBESTIMMUNG DER INTONATION EINER INTERFERENZVARIANTE DES FRAGESATZES

CHRISTINA ZACHARIAS

In den vergangenen Jahren machten u. a. Romportl und Švestková auf eine intonatorische Variante der Frage aufmerksam, die in einigen tschechischen und polnischen Dialekten vorkommt. Sie zeigt einen eigentümlichen Kadenzverlauf: Bis zur Nachakzentsilbe entspricht die Intonation einer gewöhnlichen Frageäußerung mit steigender Kadenz nach tiefliegendem Akzent. Die zweite Nachakzentsilbe wird jedoch zum Wendepunkt für die Melodieführung, die nunmehr zur Lösungstiefe strebt. Eine ähnliche Melodieform findet Weinreich im Jiddischen für die Aussage. Während er eine auf diese Weise gesprochene Aussage als „ungläubig“, „protestierend“ empfindet, sehen Romportl und Petřík die Neugier als emotionales Charakteristikum einer solchen Frageform an. Wodurch betont die „Neutralität“ dieser Gestaltung für den lachischen Dialekt.

Bei Untersuchungen zur Intonation des deutschen Fragesatzes stießen wir auf ähnliche Melodieverläufe. Es ergab sich die Frage, ob es sich auch hier nur um dialektale Besonderheiten handelte, die sich hauptsächlich auf die Gegend um Halle und Magdeburg beschränken ließen, wo die Akzente vielfach tiefer angesetzt werden als die Nachakzentsilben. Unsere Nachforschungen bestätigten das nicht. Die Wendekadenz findet beim Fragesatz auch hochsprachlich so häufig Verwendung, daß es notwendig erscheint, die Aufmerksamkeit auf diese Form zu lenken:

1. Sowohl Ergänzungsfrage als auch Entscheidungsfrage können mit Wendekadenz gesprochen werden. Der Gebrauch dieser Melodieform setzt lediglich voraus, daß in der Regel mindestens zwei Nachakzentsilben vorhanden sind, die Träger dieser Kadenz werden. Bei starker Emphase kann sie jedoch auch auftreten, wenn eine oder beide Nachakzentsilben fehlen. In diesem Fall muß die Akzent- oder Nachakzentsilbe auf langen Vokal oder einen danach folgenden Sonanten enden.

2. Die Wendekadenz vereinigt Elemente der Frage- und der Aussagemelodie. Die steigende Kadenz der sogenannten Fragemelodie, die einem tieftonigen Akzent folgt, hat eine wichtige kommunikative Eigenfunktion: Sie dient der Übertragung der psychophysischen Spannung des Sprechers auf den Hörer, der zu einer Antwortreaktion veranlaßt werden soll. Die Lösungskadenz der Aussagemelodie besitzt keine Eigenfunktion. Sie ist bloße Folgeerscheinung eines hochtonigen, starken Akzents, der die Aufmerksamkeit des Hörers auf das Detail lenkt, über das sich der Fragende informieren will. Da die Kontaktnahme dem Gebrauch der sogenannten

Fragemelodie auch dann zugrunde liegt, wenn sie an Aussagesätze gebunden ist, während das einseitige Interesse an einer Information die Verwendung der Aussagemelodie bestimmt, haben wir diese beiden grundsätzlich verschiedenen Melodieverläufe als Kontaktintonation und Informationsintonation bezeichnet und unterscheiden hinsichtlich der intonatorischen Gestaltung der Frage eine Kontaktfrage und eine Informationsfrage.

3. Bei der Wendekadenz ist die Kontaktnahme mit dem Hörer zunächst durch den steigenden Verlauf der Nachakzentsilbe angedeutet. Infolge des tiefen Akzents wird dieser Kadenzverlauf auch vom Hörer erwartet. Durch die plötzliche Wendung zur Lösungstiefe wird der bereits angebahnte Kontakt gleichsam wieder abgebrochen. Der tiefen, schwache Akzent erhält auf diese Weise eine nachträgliche, indirekte Verstärkung. Wir haben es hier offensichtlich, um mit Fonagy zu sprechen, mit einer Interferenz zweier verschiedener kommunikativer Tendenzen zu tun: einem Widerstreit zwischen Kontakt- und Informationsstreben.

4. Daraus erklärt sich die spezifische Wirkung dieses akustischen Signals auf den Hörer. Durch die plötzliche Kontaktlösung und indirekte Akzentbeschwerung entsteht der Eindruck einer emotionalen Zwiespältigkeit, als solle die kommunikative Absicht des Sprechers verschleiert oder abgeschwächt werden. Je nach dem kontextuellen Zusammenhang wird diese Frageintonation vom Hörer als lauernde, versteckt-neugierige Ausdrucksform gedeutet:

Wer \ war / denn \ das? Wo \ warst / du \ gestern?

Rhetorische Fragen gewinnen auf diese Weise die Nuance eines indirekten Vorwurfs, verhohlenen Triumphes:

Habe ich dir das nicht \ gleich / ge \ sagt?

5. Bei der Wendekadenz handelt es sich nicht um eine Doppelakzentuierung, wie man es bei dem letzten Beispiel annehmen könnte, da die zweite Nachakzentsilbe zufällig Wortakzentträger ist. In gleicher Weise werden akzentlose Silben von der Wendekadenz betroffen: Ist er benachrichtigt worden? Häufig tritt sie dort auf, wo die Frage mit einer Fragepartikel schließt:

Was \ willst / du \ denn? Was \ glaubst / du \ wohl?

6. Die Wendekadenz wird im Deutschen nicht nur mit den genannten emotionalen Schattierungen gebraucht. Ihre Kompromißstellung zwischen den beiden kommunikativen Tendenzen läßt sie verbindlicher erscheinen als die Informationsfrage, zurückhaltender als die Kontaktfrage. Daher scheint sie sich überall dort durchzusetzen, wo in der allgemeinen Verkehrssprache weder das Kontakt- noch das Informationsstreben eindeutig vorherrschen. Sie genügt dem Ausdruck des Sach-

wie des Partnerinteresses und erweckt den Eindruck zurückhaltender Verbindlichkeit, freundlicher Distanz.

7. Es läßt sich kaum beurteilen, ob diese Kadenz auf fremdsprachliche Einflüsse zurückzuführen ist oder ob sie schon immer im deutschen Sprachgebiet existierte. Ihre emotionale Bindung legt die Schlußfolgerung nahe, daß sie zum Ausdruck lauernder Neugier, versteckter Vorwürfe usw. schon immer benutzt wurde, daß ihre Verwendung zum Ausdruck gedämpfter Verbindlichkeit jedoch eine abgeleitete Funktion ist, die sie im Sinne Fonagys zur echten „melischen Metapher“ macht.

DISCUSSION

Nickel:

Es ist bekannt, daß das Geschlecht des Sprechers Einfluß auf seine Sprache hat. Das gilt für die Syntax weniger (obwohl z. B. die Häufigkeit der sogen. „progressive Form“ bei amer. Frauen höher liegt als bei Männern) als für das Vokabular, wo es ganz evident wird. Noch höher dürfte meines Erachtens aber der Unterschied innerhalb der suprasegmentalen Bereiche liegen wie z. B. hier bei der Intonation.

Zacharias:

Diese Variante ist nicht auf die Frage beschränkt, tritt aber in der Aussage seltener auf (mit zweifelnd-verwunderter Nuance, wodurch sich diese der Frage nähert. Dialektal fehlt diese Nuance). Beobachtungen und Untersuchungen zeigten, daß die These von der satzartenbestimmenden Hauptfunktion der Kadenzgestalt anfechtbar ist. Die Unterscheidung von Informations- und Kontaktintonation erfaßt auch die „Ausnahmefälle“ bei der Kadenzrealisierung der Entscheidungs- und Ergänzungsfrage und erklärt das Vorkommen der sogenannten Fragemelodie in der Aussage. Die Verwendung einer bestimmten Kadenzform wird u. a. durch das soziologische Verhältnis der Gesprächspartner bestimmt. In Situationen, in denen ein freiwilliges oder unfreiwilliges, bewußtes oder unbewußtes Abhängigkeitsverhältnis des Sprechers vom Hörer besteht, wird die steigende oder schwebende Kadenz bevorzugt. Die Rolle des soziologischen Verhältnisses für die Melodiegestaltung der Kadenz zeigt sich z. B. im Mutter-Kind-Dialog, im Vorgesetzten-Untergebenen-Verhältnis, usw.

ZUR PHONOLOGISCHEN WERTUNG DER INTONATIONSMITTEL

OSKAR ZACHER*

Die Satzintonation ist in einer bedeutenden Anzahl von Experimenten und Verallgemeinerungen dargestellt. Es hat sich Tatsachenstoff angesammelt, der einer phonologischen Systematisierung bedarf. Noch gibt es in der Wertung der Intonationsmittel keine genügende Klarheit. Wir legen eine Möglichkeit der phonologischen Wertung der Intonationsmittel vor, zu der wir auf Grund spezieller Forschungen gekommen sind.

Die phonologische Wertung der Sprachlaute beruht bekanntlich auf der Zweiteilung in Sprache und Rede. Zur Sprache zählt man die Phoneme mit ihren distinktiven Merkmalen, zur Rede die Varianten der Phoneme. Das Phonem ist selbständig, die Variante situativ bedingte. Das Wesentliche des Phonems beruht auf der Darstellung, das Wesentliche der Varianten auf Kundgabe und Appellation.

Die Wertung der Intonationsmittel kann auf ähnliche Weise erfolgen, wenn wir folgendes in Betracht ziehen.

Die Intonationsmittel bilden zusammen mit der lexisch-grammatischen Füllung des Satzes eine kompakte Einheit. Die phonologische Wertung der Intonationsmittel kann deshalb nur auf Grund der Zweiteilung in Sprache und Rede auf der ganzen syntaktischen Ebene erfolgen. Die Sprache kann auf der syntaktischen Ebene als System einer beschränkten Anzahl von „Syntaxemen“, syntaktischen Grundmodellen mit unikaler Struktur und unikaler Funktion dargestellt werden. In der Rede tritt das Syntaxem in Varianten auf, die situativ (kontextuell) bedingt sind. Wie beim Phonem beruht das Wesentliche des Syntaxems auf der Darstellung, das Wesentliche der Varianten auf Kundgabe und Appellation.

Zu den Bestandteilen des Syntaxems gehören: 1. die lexisch-grammatische Füllung in geregelter Folge, abstrahiert von der konkreten grammatischen Ausdrucksform und konkreten lexischen Bedeutungen; 2. neutraler Satzakzent; 3. Aussage-, Frage-, Befehl-, Ausrufe-, Progredienz-, Terminal-Intonem in der Ausgangsform; 4. eine verallgemeinerte syntaktische Grundfunktion (Ausdruck der Handlung, des Zustandes; Ausdruck einer Handlung, die auf ein Objekt gerichtet ist; Charakteristik des Subjekts, einer Handlung).

* Pädagogische Hochschule für Fremdsprachen Irkutsk.

Die Eigenschaften des Syntaxems treten am deutlichsten in der starken Position hervor, d. h. in neutraler gebundener Rede, die frei von intellektuellen Kontrasten und Emotionen ist. Hier treten die Syntaxeme in ihrer Grundvariante auf. In schwachen Positionen verändert sich in gewissem Maße Struktur und Inhalt des Syntaxems. Bedeutung gewinnen die situativ bedingte Wortfolge, situativ bedingte Varianten des Satzakzents (emphatischer Akzent, Neuheits- und Kontrastakzent ohne und mit Emphase), Varianten der Intoneme. Von besonderer Bedeutung ist der konkrete Inhalt des Satzes zusammen mit der syntaktischen Grundfunktion des Syntaxems.

Die Syntaxeme sind syntaktische Invarianten des Sprachsystems, in denen lexisch-grammatische Füllung, „normale“ Wortfolge, Satzakzent und Satzmelodie in ihren Grundstrukturen und -funktionen zum Ausdruck kommen. Diese Grundeigenschaften wirken in jeder Variante des Syntaxems. In der Rede treten zu den invarianten Eigenschaften variative, die situativ bedingt sind.

Die Varianten des Syntaxems sind zum Unterschied von den Varianten des Phonems relevant, da sie den Inhalt der Grundvariante konkretisieren. Die Bildung von Varianten eines Syntaxems erfolgt nach bestimmten Gesetzmäßigkeiten, die in der Rede wirken. Auf syntaktischer Ebene sind daher zu unterscheiden: 1. die ständig wirkende Relevanz der Syntaxeme des Sprachsystems, 2. die situativ bedingte Relevanz der Varianten des Syntaxems in der Rede.

Insofern der Systemcharakter nicht nur in den invarianten Eigenschaften des Syntaxems (und Phonems) zum Ausdruck kommt, sondern auch in den variativen Eigenschaften der Modifikationen des Syntaxems (und Phonems), sind Sprachsystem und Redesystem zu unterscheiden, die zwei Stufen einer dialektischen Einheit — der Sprache — darstellen. Die gesamten Intonationsmittel der Sprache lassen sich somit auf Grund von Unterscheidung syntaktischer sprachsystemlicher Grundrelevanz und variativer Relevanz des Redesystems systematisieren.

ПРОБЛЕМА РАСПОЗНАВАНИЯ РЕЧЕВЫХ СИГНАЛОВ КАК СЛОЖНАЯ СИСТЕМА

Н. Г. ЗАГОРУЙКО*

Речевой сигнал, переходя в процессе восприятия с одного иерархического уровня на другой, претерпевает многократную перекодировку. Теория сложных систем (1) позволяет найти ответ на следующий важный вопрос: *В чем смысл иерархической структуры восприятия речи?*

Ответ состоит в том, что: 1. в результате перекодировки удастся согласовать характеристики сигнала на выходе одной ступени с характеристиками простого классификатора следующей ступени; 2. при многоуровневой системе требуется небольшая оперативная память; 3. на каждом уровне появляется возможность использовать дополнительную информацию о речи и языке, которая может храниться в долговременной памяти.

В результате, с помощью простых структур при малых затратах оперативной памяти удастся достигать надежного распознавания устного сообщения.

За счет чего можно упростить структуру каждого уровня?

Нами обнаружено (2), что в процессе принятия решения в многомерном пространстве признаков человек использует самый простой тип решающих границ-гиперплоскости, перпендикулярные осям координат. Это накладывает ограничение на характер распределения распознаваемых элементов-распределения должны обладать малой дисперсией. Уменьшение дисперсии может достигаться путем различных видов нормализации — по громкости, темпу, особенностям дикции и т. д. Экспериментальное исследование алгоритма подстройки эталонов под диктора (3) показала явную целесообразность этого вида адаптации.

В очень сильной степени сложность процедуры распознавания зависит от количества распознаваемых объектов. Предварительное сокращение промежуточных алфавитов может быть получено с помощью алгоритмов таксономии (4, 5). Большой выигрыш во времени дает сокращение альтер-

* Институт математики Сибирского отделения АН СССР г. Новосибирск.

натив за счет априорной информации, а так же за счет информации, поступающей в процессе распознавания. Алгоритмы, реализующие такую возможность, имитируют хорошо известный в психологии феномен „установки“ и предложенный Л. А. Чистович метод „вычеркивания“ (6).

Рассмотрение известных фактов о восприятии речи человеком с позиций теории сложных систем приводит нас к варианту структуры многоуровневого распознающего устройства, представленного на рис. 1. Основные элементы этой схемы обсуждались с Бондарко Л. В., Молчановым А. П., Кожевниковым В. А. и особенно часто с Чистович Л. А. Поэтому, заключительную часть доклада можно считать результатом нашей совместной работы.

Нам представляется, что речевой сигнал $f(t)$ преобразуется в самом начале в некоторое довольно полное описание в пространстве (X_1) частота-время. На участке длительностью в несколько сотен миллисекунд определяются некоторые признаки (S_1) типа статических и динамических характеристик формант, характеристик шумовой части спектра и т. д. Разумеется, надежность распознавания (P_1) этих признаков может оказаться недостаточно высокой, даже после использования всех возможностей классификатора D_1 . Повысить эту надежность можно за счет использования блоком H_1 информации о физических законах речеобразования такого типа: частота основного тона не может меняться быстрее некоторой величины; одновременное существование таких-то признаков невозможно; вслед за такой комбинации признаков наиболее вероятно появление таких-то признаков и т. д. Использование информации подобного рода продолжается до тех пор, пока вероятность некоторого варианта признаков не станет больше вероятности других вариантов на некоторую пороговую величину ΔP_1 . Этот вариант поступает на вход следующей ступени преобразования, где последовательность таких признаков образует пространство описания X_2 .

Если разница вероятностей меньше ΔP_1 , то запоминается несколько вариантов признаков S_1 .

На втором уровне происходит распознавание фонем (S_2) , вернее их контекстуальных вариантов (звукотипов). С этой целью классификатор D_2 использует информацию на участке типа открытого слога, границы которого определяются сегментатором C_2 . Для сокращения числа возможных вариантов некоторой фонемы используется информация, содержащаяся в описании X_2 , а затем, если необходимо, и информация о структуре последовательности фонем. С этой целью блоком H_2 формируются последовательности наиболее вероятных вариантов фонем и, с учётом всех этих постериорных и априорных сведений, выбирается наиболее вероятная последовательность (7). Если разность вероятности этой избранной последовательности и любой другой превышает некоторый порог ΔP_2 , то фонемный код слога передается на вход следующего блока. В противном случае категорического решения не принимается и запоминаются коды фонем S_2 несколь-

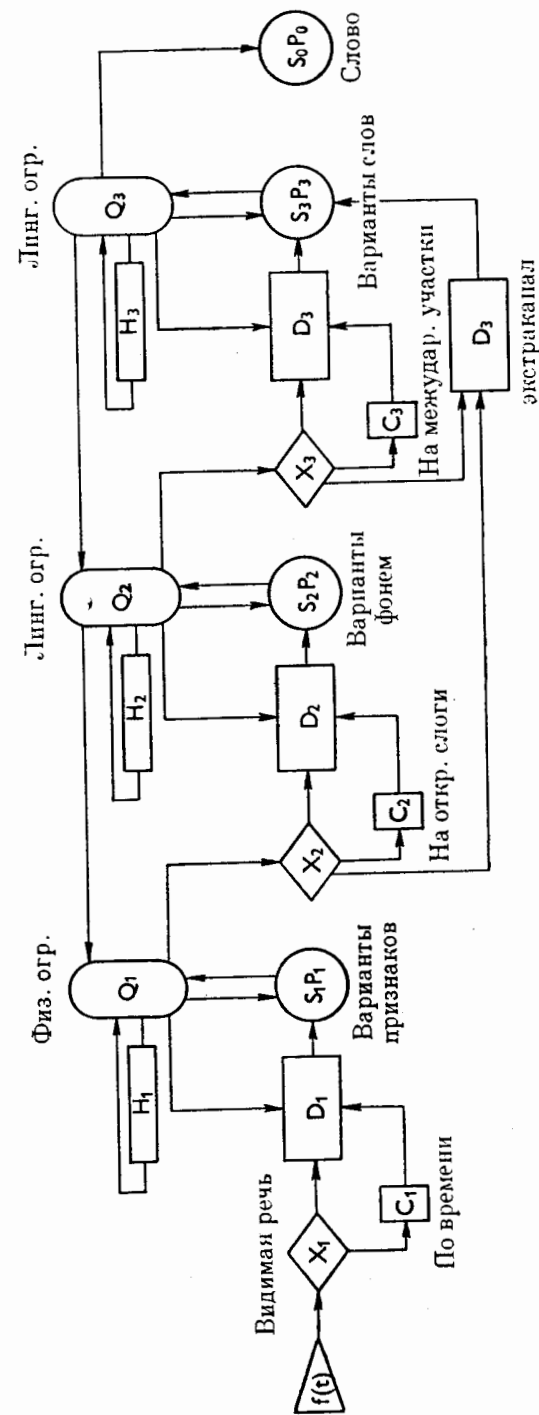


Рис. 1.

ких (наиболее вероятных) слогов. Если этих вариантов слишком много, то можно повторить процедуру, вызвав на вход блока другой вариант признаков X_2 .

Для распознавания слов из словаря S_0 пространство X_3 кроме кодов фонем должно содержать информацию и об ударениях. Сегментатор C_3 осуществляет членение речевого потока на участки от ударения до ударения. В двух таких соседних участках содержится, как минимум, одно слово словаря S_0 . Поиск нужного слова и одновременное определение его границ может осуществляться с помощью алгоритма Лисенко (8). На этом этапе, как и раньше, при выборе решения могут использоваться (блоком H_3) дополнительные априорные сведения об элементах словаря и, если окажется необходимым, осуществляться вызов (по линии Q_3-Q_2) на вход (X_3) нескольких вариантов фонемных последовательностей.

Наряду с этим должен существовать механизм (D_3) распознавания слов по неполной информации — например, только по началу или только по некоторому ограниченному набору признаков на определенном участке слова.

Рассмотрение такой модели позволяет наметить цели конкретных исследований, необходимых для конструирования искусственных систем распознавания речи.

ЛИТЕРАТУРА

1. Гуд, Г. Х., Макол, Р. Э.: Системотехника. Введение в проектирование больших систем. Перевод с англ. Изд-во „Советское Радио“, Москва, 1962.
2. Загоруйко, Н. Г.: Какими решающими функциями пользуется человек? Сб. Тр. ИМ СО АН СССР „Вычислительные системы“ вып. 28, Новосибирск, 1967.
3. Загоруйко, Н. Г., Лозовский, В. С.: Подстройка, под диктора при распознавании ограниченного набора устных команд. Сб. Тр. ИМ СО АН СССР „Вычислительные системы“ вып. 28, Новосибирск, 1967.
4. Елкина, В. Н., Загоруйко, Н. Г.: Об алфавите объектов распознавания. Сб. Тр. ИМ СО АН СССР „Вычислительные системы“ вып. 22, Новосибирск, 1966.
5. Загоруйко, Н. Г., Елкина, В. Н.: Алфавит с минимальной избыточностью. Сб. Тр. ИМ СО АН СССР „Вычислительные системы“ вып. 28, Новосибирск, 1967.
6. Загоруйко, Н. Г.: Комбинированный метод принятия решений. Сб. Тр. ИМ СО АН СССР „Вычислительные системы“ вып. 22, Новосибирск, 1966.
7. Волошин, Г. Я.: Об использовании языковой избыточности для повышения надежности автоматического распознавания речевых сигналов. Сб. тр. ИМ СО АН СССР „Вычислительные системы“, вып. 28, Новосибирск, 1967.
8. Лисенко, Д. М.: Выделение и морфологический анализ слов в речевом потоке. Диссертация. Ленинград, 1966.

DISCUSSION

Райнов:

Вопрос: Каким образом осуществляется вычеркивание элементов, которые, как вы показали, выполняют функцию избыточности?

И второй вопрос: Какой процент надежности обеспечивает это устройство для распознавания?

Загоруйко:

К Райнову: 1. Алгоритмы сокращения избыточности алфавита описаны в работах Загоруйко Н. Г. и Елкиной В. Н. в трудах Института математики СО АН СССР (г. Новосибирск) „Вычислительные системы“ № 22, 1967 г.

2. Языковая избыточность сильно повышает надежность распознавания фонем. Так, в опыте со 125 словами информация о правилах сочетания двух фонем повысила процент правильно распознанных фонем с 15 до 94. Этот алгоритм описан в работе Волошина Г. Я. в трудах Института математики СО АН СССР (г. Новосибирск) „Вычислительные системы“, № 28, 1967 г.

К Фанту: Мы считаем, что элементами распознавания на первом уровне должны быть статические и динамические характеристики формант—частота, ширина, скорость движения. Кроме этого должны распознаваться признаки источника-шума, голоса и т. д.

QUELQUES REMARQUES CONCERNANT LA DURÉE RELATIVE DES VOYELLES DU FRANÇAIS CONTEMPORAIN

MARIE ZICHOVÁ

Nous voulons nous limiter à quelques remarques concernant la longueur relative des voyelles françaises dans les styles différents.

Pour pouvoir comparer les longueurs des voyelles, nous avons examiné quatre échantillons du langage parlé.

Le premier groupe représente la déclamation de vers (deux poèmes de P. Verlaine, interprétés par deux artistes);

le deuxième groupe la lecture de la prose (le texte lu par deux interprètes différents);

le troisième groupe une discussion concernant les problèmes du théâtre contemporain;

le quatrième groupe une discussion spontanée d'étudiants parisiens.

Dans le premier groupe nous avons fait l'analyse de deux poèmes en examinant les valeurs minima et maxima des voyelles accentuées et inaccentuées. Dans les trois autres groupes nous nous sommes contentés du choix à peu près au hasard de 150 voyelles pour chaque groupe — et nous avons comparé de nouveau les valeurs minima et maxima des voyelles accentuées et inaccentuées — ainsi que des voyelles longues et brèves. En mesurant la longueur des voyelles on a confronté l'enregistrement sonore avec les valeurs objectives des spectrogrammes.

1° Premier groupe:

Il est bien naturel que l'interprétation des vers de P. Verlaine soit assez „artificielle“ en comparaison avec la prononciation courante du langage parlé. (Le style de déclamation de nos deux exemples est tout à fait moderne et ne tombe jamais dans l'expression pathétique des contrastes exagérés.) Dans ce style, nous pouvons distinguer deux groupes de voyelles: les voyelles *accentuées* et les voyelles *inaccentuées*, dans ce sens que les voyelles inaccentuées se trouvent dans la limite de 60—125 msec, tandis que les voyelles accentuées se trouvent dans la limite de 130—290 msec. On peut remarquer que la limite maxima des voyelles inaccentuées représente à peu près la limite inférieure des voyelles accentuées. Dans quelques cas, les valeurs des voyelles longues et brèves sont un peu différentes, — mais, d'autre part, on trouve beaucoup

* Institut de Phonétique, Université Charles, Prague.

de voyelles dites longues qui sont de la même longueur que les voyelles brèves à condition que la rapidité du débit, le tempo, reste toujours à peu près le même.

2° Deuxième groupe:

Lecture de la prose est très naturelle et on pourrait bien la comparer au style élevé des paroles publiques. La limite inférieure de 150 cas examinés est à peu près 40 msec (surtout pour les voyelles inaccentuées), tandis que la limite supérieure des voyelles brèves dépasse très souvent la valeur des voyelles longues (190, 200, 210 msec) — même si le tempo ne change pas.

Exemples:

Dans le mot *grandir* [ã] inaccentué dure 50 msec, tandis que dans le mot *encore* [ã] inaccentué dure 120 msec — et le [ɔ:] suivant seulement 90 msec. Les [e] finals des mots *crever* et *expliquer* sont aussi bien différents: 200 et 70 msec. Dans le mot *mensonge* les deux voyelles nasales: [ã] bref et inaccentué — [õ:] long et accentué sont de la même longueur 150 msec.

3° Troisième groupe:

Peut-être, est-il un peu surprenant que même si la limite inférieure des voyelles de la discussion spontanée reste à peu près 40—50 msec, la limite supérieure ne dépasse que très rarement 150 msec pour les voyelles brèves et même 190 msec pour les v. longues.

Exemples:

Le [o:] long dans le mot *chose* dure dans deux cas 70 msec et une fois 100 msec; le [oe:] long du mot *spectateur* a toujours la valeur de 70 msec; le [a:] long dans le mot *de ma part* dure 90 msec; le [i:] long dans le mot *il faut dire* dure 70 msec et le [ɔ:] long dans le mot *tout d'abord* dure 80 msec — tandis que le [ɛ] bref des mots *culturelle* et *professionnelle* dure 150 et 110 msec.

4° Quatrième groupe:

La limite inférieure reste à peu près la même (40—50 msec), mais la limite supérieure est bien variable: 70, 80, 90, mais aussi 200 et 210 msec pour les voyelles brèves ainsi que pour les voyelles longues. Dans la discussion des étudiants nous trouvons des contrastes encore beaucoup plus frappants.

Exemples:

Les [e] et [ɛ:] du mot *tu préfères* sont exactement de la même longueur: 70 msec et nous trouvons p.ex. la même longueur de [ɛ:] dans le mot *dernière*. Dans la phrase: *Ma chemise sera sèche*, le [i:] long dure 80 msec, le [a] bref du mot *sera* dure 80 msec et le [ɛ] du mot *sèche* dure 110 msec. Un exemple bien frappant se présente dans les mots *ma fenêtre* — où le [a] inaccentué et bref dure 70 msec, tandis que le [ɛ:] accentué ne dure que 60 msec.

Conclusion:

En examinant les quatre échantillons différents — pour l'instant encore d'une façon incomplète — nous pouvons dire le suivant:

Plus le style est soutenu, plus on distingue les voyelles accentuées et inaccentuées et même, quelque fois, les voyelles longues et les voyelles brèves — tandis que dans un style du français parlé spontanément les contrastes des longueurs ne se montrent pas toujours d'une façon très nette. Ce qui influence toujours la longueur des voyelles — longues et brèves — d'une façon particulière, c'est l'accent d'insistance — qui peut allonger dans n'importe quel style une voyelle normalement brève — comme p.ex. le [i] dans le mot *Paris* (200—250 msec); cela s'applique même à E-caduc dans une expression comme: „vous voyez le spectateur...“ où le [ə] dure 90 msec — et normalement 40 msec. En tout cas, la longueur ne se présente pas comme un facteur phonologique, mais purement phonétique.

DISCUSSION

Gsell:

Il est bien évident que la distinction entre longues et brèves en français est un fait de norme et non pas de système. Il n'y a pas d'opposition de quantité si bien que le recouplement ne doit pas nous étonner. Une autre cause c'est le fait qu'on n'ait pas tenu compte des durées spécifiques: une voyelle nasale même inaccentuée est toujours plus longue qu'une autre voyelle même „longue“ accentuée, d'où la nécessité de tenir compte de facteurs de pondération.

Mme Gsell:

Il me semble difficile de différencier brutalement les styles par les durées. En effet, un style très soutenu, d'un niveau poétique indiscutable, peut être dit très vite. De la prose peut se dilater. Le temps d'un texte varie selon les habitudes d'un individu, ses sentiments du moment, le contenu particulier du message. De toutes façons, on ne peut établir de système sur un échantillonnage aussi restreint.

Witloch:

1. La longueur de voyelles en français n'a pas de valeur phonologique — voilà l'essentiel, à notre avis.

2. Je mentionne les travaux de J. Chlumský concernant la longueur des voyelles tchèques: celles courtes sont prononcées „staccato“, celles longues „legato“. Leur longueur réelle n'est pas tellement importante.

О ВОСПРИЯТИИ РЕЧЕВОЙ ИНТОНАЦИИ

И. А. ЗИМНЯЯ*

Изучению механизма восприятия речевой интонации в целом и одного из ее ведущих компонентов частоты основного тона, в последнее время посвящено достаточно много работ (В. А. Артемов, Л. А. Вербицкая, Ю. Бэнинг и Ван-Скунефельд, Н. И. Жинкин, И. А. Зимняя, П. В. Исаченко, Н. А. Любимова, Ф. Либерман, Е. Улдалл и др.). Усилия исследователей направлены на выявление либо отдельных параметров (Е. Улдалл и др.), либо структуры параметров (В. А. Артемов и др.), которые могут определить собой опознавание того или иного коммуникативного типа интонации (напр. вопроса или повествования).

Целью данной работы является изучение зависимости опознавания коммуникативного типа интонации 1. от положения ударного гласного в слоговой структуре слова — предложения и 2. от направления движения основного тона во всей структуре. Теоретическими предпосылками данного исследования были следующие:

1. В памяти человека, владеющего данным языком, имеется алфавит интоном, которые носят зонный характер возможных физических реализаций данного языкового эталона.

2. Распознавание и отождествление предъявляемой интонационной реализации происходит в результате ее сопоставления с зонным эталоном на основе принятия решения о тональном завершении структуры и локализации определенного изменения направления тона в ударном гласном.

3. Всякое изменение привычного соотношения слоговой структуры слова и ее интонационного оформления может вызвать затруднение в опознавании коммуникативного типа интонации и возможность оценки данной реализации только в эмоциональном или волевом плане.

Исследование проводилось на материале слов — предложений типа (С)ГСГ(С) и (С)ГСГ(С). Были подобраны слова „перевертыши“, типа шалаш, довод. Слова произносились с интонацией вопроса и ответа повествования. Всего было записано 18 слов в произнесении одного диктора.

* Кафедра психологии 1 МГУИИИЯ, Москва.

Вслед за тем 13 аудиторов (из них 3 высококвалифицированных фонетиста) прослушивали обратное по ходу записи воспроизведение речевого сигнала, регистрируя то, что выражается интонацией — вопрос, повествование, какое-либо воздействие или эмоциональное состояние.

Результаты аудиторского прослушивания были соотнесены с результатами интонационного анализа изучаемых структур.

Эксперимент показал, что коммуникативный тип (вопрос или повествование) уверенно опознается при совпадении слоговой структуры и ее интонационного оформления характерного для данного языка. — Для опознавания интонации повествования первой слог обеих структур должен быть выше по частоте основного тона второго слога, вне зависимости от того, является ли он ударным или нет.

1. При несовпадении структуры и ее интонационного оформления, решающим является не только финальное завершение тона, но и то, реализуется ли оно в ударном или безударном слоге.

2. Это оформление возникает в результате перевертывания обеих структур, первоначально произнесенных с интонацией повествования.

3. В этом случае первый слог как ударный, так и неударный оказывается ниже второго при общем восходящем движении тона.

Если второй слог был неударным, то только при интервале падения тона $\approx 1\frac{1}{2}$ октавы, структура опознавалась как повествование. При интервале падения, стремящемся к секунде, в 45 % случаев эта структура опознавалась как вопрос и в 45 % как повествование. При интервале подъема в приму структура в 63 % случаев была опознана как вопросительная.

При столь же малых интервалах падения в ударном слоге в 85 % случаев структура опознавалась как повествовательная.

Обратное прослушивание и опознавание структуры СГСГ, произнесенной с интонацией вопроса, подтвердило гипотетические предпосылки работы. Получившаяся в результате обратного воспроизведения структура отличается от свойственной языку структуры только тем, что подъем $\approx 1\frac{1}{2}$ октавы реализуется не в ударном, а в безударном гласном. И несмотря на то, что ударный гласный характеризуется глубоким, завершающим структуру падением частоты основного тона, она опознается в 45 % случаев как вопрос и в 45 % как повествование. При опознавании происходит как бы „сбой“ между движением частоты основного тона и его локализацией в структуре.

Существенны результаты опознавания перевернутой структуры СГСГ, произнесенной с вопросом. Полученная в результате структура очень близка по движению частоты основного тона к структуре СГСГ к повествованию. Но так как глубина интервала падения $\approx 1\frac{1}{2}$ октавы не укладывается в зону эталона чистого коммуникативного типа повествования и распространяется на зоны эмоциональных интоном — удивления, страха

и т. д., то полученная структура опознается в эмоционально-волевом тоне.

Анализ полученных данных дает основание предположить, что не только финальное тональное завершение релевантно для опознавания оппозиции вопрос—повествования (Е. Улдалл), но и соотношение его со слоговой структурой слова. Подтверждается также положение о вероятностном характере опознавания в процессе соотнесения с зонным эталоном памяти.

ФОНЕМА И ВОСПРИЯТИЕ

Л. Р. ЗИНДЕР*

Рассмотрение фонемы как члена противопоставления привело Трубецкого к различению релевантных (по современной терминологии — дифференциальных) и иррелевантных признаков. Так как для выполнения дистриктивной функции, т. е. для различения значимых единиц языка (морфем, слов) фонема должна лишь отличаться от всех прочих фонем данного языка, то с точки зрения этой функции фонологическое содержание фонемы можно считать исчерпывающимися дифференциальными признаками. Фонема в таком случае может быть определена как сумма или пучок дифференциальных признаков. Последние в такой трактовке точнее называть вслед за С. К. Шаумяном дифференциальными *элементами*.

При таком понимании вещей в качестве дистриктора, строго говоря, выступает не фонема, а дифференциальный элемент, так как именно он различает фонемы, а через них и более высокие языковые единицы. Вместе с тем дифференциальный элемент, будучи структурной, а не линейной величиной, проявляется только через фонему, или в речи — через варианты фонемы. Вследствие этого фонеме и приписывают дистриктивную функцию.

Установление системы языка как таковой не требует непосредственного обращения к психике его носителей. Однако язык, являющийся средством общения между людьми, должен рассматриваться в лингвистике не как статистическая, а как функционирующая система, которая выявляется в речи. Это означает, что лингвисты не могут обходить вопроса о восприятии речи носителями соответствующего языка и о значении для восприятия отдельных языковых единиц.

Говоря о дистриктивной (смыслоразделительной) функции фонемы, Л. В. Щерба и Н. С. Трубецкой подчеркивали, что эта функция обнаруживается не только в том, что при замене фонемы одно слово (или фонема) может заменяться другим, но и в том, что замена фонемы может разрушить слово, превратить его в бессмысленное звуко сочетание (точнее — фономосочетание). В последнем случае говорить о дистриктивной функции не-

* ЛЭФ Ленинградского университета — Ленинград.

уместно; здесь дело идет о том, что слово не может существовать без соответствующего состава фонем и определенного порядка следования их. В аспекте восприятия это означает, что слово в таких условиях не может быть опознано, или же что опознание его, по крайней мере, становится затруднительным. Фонема, следовательно, выступает здесь в конститутивной, или иначе — в словоопознавательной функции.

Различие между вариантами одной фонемы не может быть использовано для различения слов или морфем, однако замена одного варианта другим затруднит восприятие слова. Показывая на примере слов *мел* и *мель*, что разные „e“ являются в русском языке вариантами одной фонемы, а согласные „л“ и „ль“ представляют две фонемы. С. И. Бернштейн замечает: „Разница между этими согласными достаточна для различения пары приведенных слов, но говорящие, различая слова *мел* и *мель*, опираются на различие общего облика слов, зависящее в равной мере от разницы в согласных и в гласных.“ (Фонема. Большая советская энциклопедия, том 58, 1936.) Экспериментально-фонетические исследования, проведенные в последнее время в лаборатории Ленинградского университета, показали, что правильное восприятие палатализованных согласных в значительной степени зависит, а иногда и целиком определяется соответствующими акустическими признаками следующих за ними гласных (Л. Р. Зиндер, Л. В. Бондарко, Л. А. Вербницкая. Акустическая характеристика различия твердых и мягких согласных в русском языке. „Вопросы фонетики“, 1964); при этом эти признаки являются irrelevantными, так как они не используются в русском языке для противопоставления гласных фонем.

Все сказанное заставляет пересмотреть вопрос о фонологическом содержании фонемы. С точки зрения восприятия, опознавания значимых единиц языка (слов, морфем) физиологическое содержание фонемы не может исчерпываться дифференциальными элементами, в него входят и irrelevantные признаки. Так, переднеязычность не является дифференциальным элементом негубного носового согласного в русском языке, так как в нем нет заднеязычного носового, которому он бы был противопоставлен фонологически. Тем не менее, замена переднеязычного согласного заднеязычным в таких словах, как например, нос или сон „разрушит“ эти слова так же, как если бы в них „н“ был заменен согласным „л“. Разумеется, для носителей русского языка такого рода замены не равнозначны: *нос* непонятно, но звучит по-русски, *ос* же воспринимается как нечто чуждое русскому языку, (ср. Л. Р. Зиндер. Общая фонетика. Л., 1960, стр. 40). В других случаях, а именно, перед заднеязычными согласными дело обстоит несколько по-иному: слово *анкета*, например, произнесенное с заднеязычным носовым, опознается правильно, но при этом ощущается „иностраный акцент“, в той или иной степени затрудняющий восприятие речи.

Таким образом, можно сказать, что пренебрежение irrelevantными

признаками ведет к разрушению слова, или же осложняет его опознавание. Следовательно, определение фонемы как пучка дифференциальных элементов, по крайней мере, с точки зрения восприятия не может быть признано удовлетворительным. Этими ставится под сомнение и то первостепенное значение, которое придается в определении фонемы дистриктивной функции ее. Знаки языка могут различаться и в том случае, когда план выражения их совпадает. Билатеральная природа делает это совершенно очевидным: для того, чтобы два знака, из которых каждый обладает двумя сторонами, различались между собой, достаточно несовпадение одной из этих сторон.

Резюмируя все сказанное, следует признать, что основной функцией фонемы является конституирующая или опознавательная (словоопознавательная или морфемоопознавательная), дистриктивная же функция будет производной, вторичной.

DISCUSSION

Artemov:

Я не усматриваю противоречия между новитями: фонема как пучок дифференциальных признаков и восприятие фонем как совокупность перцепируемых качеств. Человек воспринимает звуковой образ слова, а не сумму фонем.

Wiede:

Я думаю, что распознавание слова осложняется только тогда, когда аллофоны произносятся неправильно. Но оно все-таки еще существует. Так, например, русские произносят слово „dich“ с твердым [x], но немцы их понимают, а немцы произносят слово „bank“ с [ɔ] и их также понимают. Мне кажется, что общение нарушено только тогда, когда неправильно произносятся фонемы.

ВРЕМЕННЫЕ ХАРАКТЕРИСТИКИ РУССКОЙ РЕЧИ

Л. В. ЗЛАТУСТА*

1. Исследования длительности звуков русской речи показали, что различия по абсолютному времени звучания как согласных, так и гласных сохраняются только в тождественных позициях. В потоке речи наблюдается „пересечение“ длительности почти всех типов согласных и гласных звуков.

2. В связи с этим целесообразно рассматривать относительную длительность звуков в пределах определенной речевой единицы.

За единицу, в пределах которой рассматривается относительная длительность звуков речи, принята нами ритмическая структура, представляющая собой знаменательное слово или сочетание служебного и знаменательного слов, объединенных одним словесным ударением. (В отдельных случаях в состав ритмической структуры может входить два или несколько знаменательных слов.)

Закономерная последовательность звуков, составляющих ритмическую структуру, организуется словесным ударением.

Словесное ударение не обязательно связано с конкретным звуком или слогом, наделенным специфическими акустикофизиологическими качествами ударности; место ударения определяется слушающим благодаря закономерным отношениям звуковых единиц внутри ритмической структуры. Так, в структуре $Г С \acute{Г} С Г$ первый гласный и второй могут быть совершенно одинаковы по длительности, спектральному составу, размеру ротовой щели и т. д., но смещения со структурой $\acute{Г} С Г С Г$ не произойдет, т. к. во втором типе структуры никогда не наблюдается совпадения данных по перечисленным параметрам для двух сравниваемых звуков.

Для структуры $Г С Г С \acute{Г}$ также характерно совпадение начального гласного с ударным по указанным параметрам.

На рис. 1 дан график, иллюстрирующий результат спектрального анализа гласного "а" под ударением и во II предуд. слоге. График показывает разности суммарной интенсивности полос, на которые делится весь спектр

* Московский государственный университет.

речи. Полосы выбраны на основании предварительных исследований спектральных картин. По оси абсцисс отложены разности уровней в дБ, по оси ординат — число случаев значений в процентах для каждой полосы.

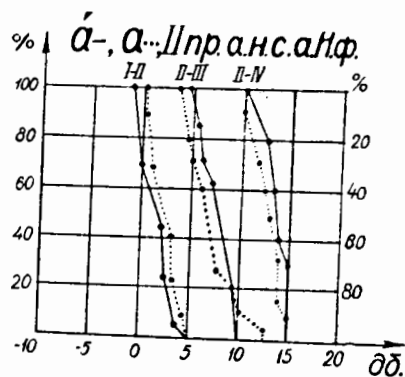


Рис. 1.

Рис. 2 — спектрограмма "а" и рис. 3 — спектрограмма "а" 1 предуд. слога показывают индивидуальные спектры гласных, на основании которых набирались статистические данные для построения интегральных кривых, представленных рис. 1.

На рис. 4 дана фотография, показывающая раствор ртовой щели при произнесении структуры ГСГСГ'СГС, где максимальный раствор наблюдается на II предударном гласном.¹

Наиболее типичное распределение значений ширины раскрытия ртовой щели таково, что в структуре ГСГСГ' или ГСГСГ'С максимум наблюдается на II предуд., некоторое падение на 1 предуд. и возрастание на ударном.

В структуре ГСГ'СГ1 предударный и ударный обладают близкими характеристиками. Всегда большим раскрытием ртовой щели отмечен ударный гласный в структуре Г'СГСГ. Условно эти отношения можно показать следующей схемой:

ГСГСГ' — — — — —
 ГСГСГ — — — — —
 Г'СГСГ — — — — —

Набор классов ритмических структур в речи ограничен. В основу классификации ритмических структур на классы положено два фактора: место

¹) Описание прибора помещено в книге Л. А. Чистович и др. Речь. Артикуляция и восприятие. Изд. Наука. Москва—Ленинград, 1965 г.

словесного ударения и последовательность типов звуков в ритмической структуре.

Эксперимент по исследованию деформации ритмической структуры в зависимости от условий контекста выявил устойчивость ритмической структуры.

Исследование количества ритмических структур для русской речи (на основании аудиторского анализа структур с искажением спектра, но сохранением огибающей интенсивности и частоты основного тона) дало приблизительную цифру 15.

Проделанная работа и имеющиеся в литературе вопросы данные позволяют выдвинуть гипотезу о наличии минимум трех параллельно действующих уровней при восприятии речи.

1. Звуковысотный уровень, „заведующий“ набором интонационных контуров, т. е. наиболее общими смысловыми характеристиками. Их количество ограничено — примерно 6 для русской речи.

2. Уровень ритмических структур, примерно 16.

3. Уровень опознавания последовательностей звуков, но лишь в их „сильных“ т. е. информативных вариантах, т. к. слабые варианты — материал, сохраняющий тип ритмической структуры.

DISCUSSION

I. Mahnken:

Ich möchte zunächst betonen, daß ich mit der von Ihnen vorgelegten Konzeption völlig übereinstimme. Ich möchte nur eine Frage stellen: Bekanntlich werden im Russ. am Ende einer Aussage oder bestimmter Aussageabschnitte auch die unbetonten Silben (bzw. Vokale) sehr stark gedehnt, so daß sie dann oft länger sind als die Akzentsilbe. Sie betrachten die Dauer als relevantes Merkmal der Akzentuiertheit im Russischen. Wie erklären Sie die Tatsache, daß in den erwähnten Fällen trotzdem die Akzentuiertheit der Akzentsilbe richtig erkannt wird?

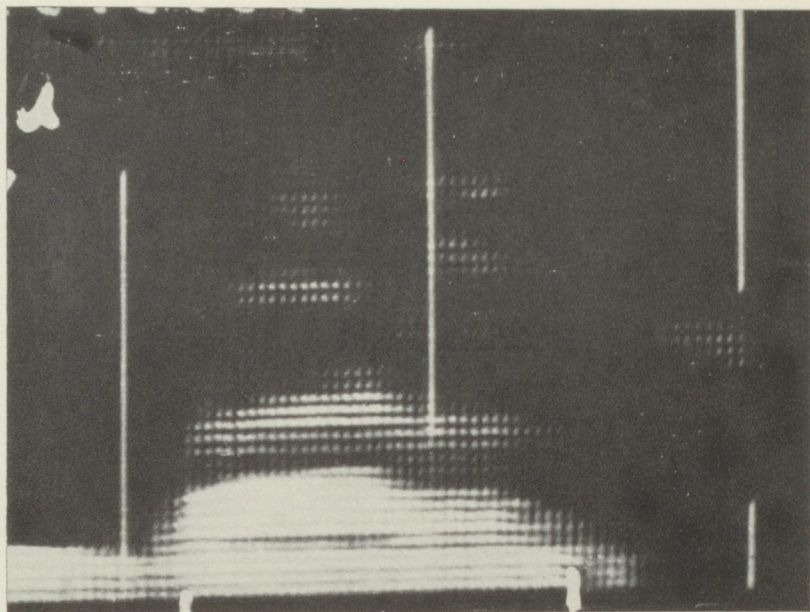


Рис. 2.

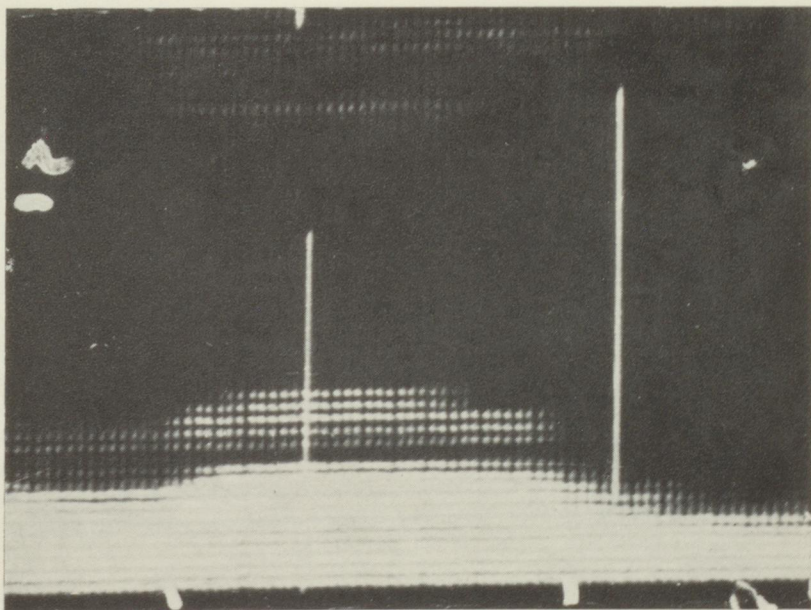


Рис. 3.

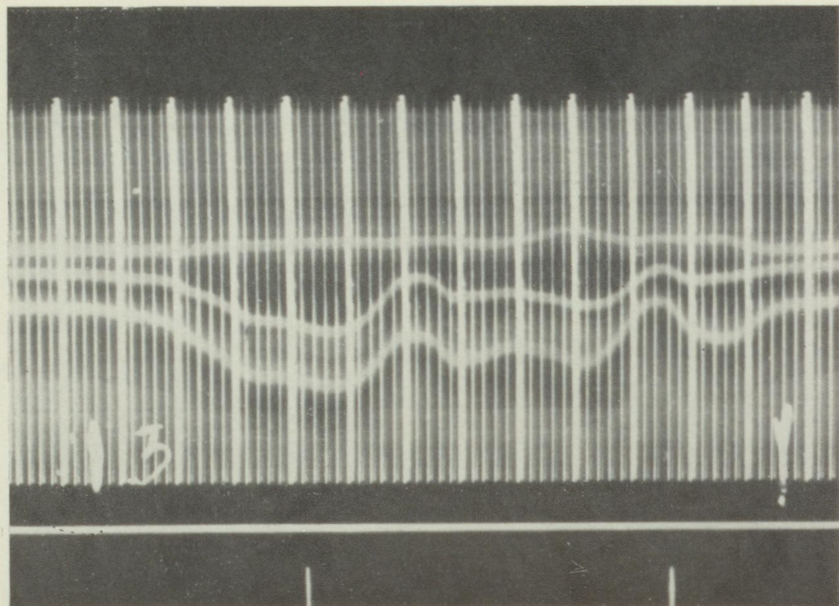


Рис. 4.

LA RÉUNION DU CONSEIL PERMANENT POUR L'ORGANISATION DES CONGRÈS INTERNATIONAUX DES SCIENCES PHONÉTIQUES

La réunion du Conseil Permanent a eu lieu le 9 Septembre 1967 avec la participation des membres suivants:

Président Prof. D. B. Fry
Vice-Président Prof. B. Hála
Membres: Prof. V. A. Artemov
 P. Delattre
 G. Fant
 E. Fischer-Jørgensen
 M. Kloster Jensen
 B. Malmberg
 A. Martinet
 M. Onishi
 W. Pée
 A. Rigault
 A. Sovijärvi
 F. Trojan

Ordre du jour:

1. Nouveaux membres:

Ont été élus:

Prof. L. K. Pike (USA) à la place du Prof. G. E. Peterson (USA), décédé;

Prof. G. Ungeheuer (RFA), à la place du Prof. E. Zwirner (RFA) qui renonce à la fonction de membre actif;

Prof. E. Zwirner (RFA) prend la place du Prof. M. P. Fouché (France), décédé, membre honoraire.

2. Lieu du prochain Congrès:

La proposition du † Prof. Peterson d'organiser le 7e Congrès aux États Unis a été reconsidérée, ainsi que celle du Prof. Rigault qui propose la ville de Montréal, au Canada; faute d'une approbation officielle des organismes canadiens compétents, la décision finale a été ajournée.

En cas qu'une proposition supplémentaire des États-Unis parviendrait au Conseil, le Président Fry déciderait définitivement entre le Canada et les États-Unis; néanmoins, la proposition canadienne est au premier rang.

3. Conformément à la proposition du Professeur Rigault, le prochain congrès aura lieu en 1971. Toutes les résolutions faites pendant la réunion du Conseil Permanent ont été adoptées à l'unanimité.

R. Gsell, Secrétaire

REPORT OF THE GENERAL MEETING OF THE INTERNATIONAL SOCIETY OF PHONETIC SCIENCES

Election of President: Professor Malmberg.

Election of Committee to decide on appointment of Subsecretaries and Delegates:

Malmberg, Winter (Kiel, Germany), Gimson (London), Kloster Jensen.

Membership fee: \$ 2 (or counter-value) a year.

A Walk Over of the existing Board of Directors was decided on and Professor Avanesov (Moscow) was elected to join the Board as a Vice-President. The Board of Directors thus consists of the following persons: President: Malmberg, Vice-Presidents: Zwirner, Artemov, Jakobson, Onishi, Avanesov, Secretary: Kloster Jensen, Treasurer: Petersson.

The preliminary draft of By-Laws was accepted provisionally in a partly changed form.

Further decisions:

The Board was requested to consider the techniques of election and to propose an appropriate procedure at the next general meeting.

The General Meeting further considers it a task of the officers to have relationships established between subdivisions and local corporations of interest to phoneticians.

End of Report

According to the By-Laws, Article Six, Section One, the Board of Directors have elected the following Officers besides the President (elected by the General Meeting): Zwirner, Kloster Jensen, Petersson.

BY-LAWS PROVISIONALLY ACCEPTED AT THE
GENERAL MEETING OF THE ISPhS,
SEPTEMBER 9th, 1967

Article One

ORGANIZATION

1. The name of this organization shall be INTERNATIONAL SOCIETY OF PHONETIC SCIENCES, shortened: ISPhS.
2. The organization may at its pleasure by a vote of the membership body change its name.
3. The Board of Directors shall consist of a President, one or more Vice-Presidents, a Secretary and a Treasurer.
4. Subdivisions shall be established in single countries, and membership fees shall be paid to the subdivisions in the respective countries. Such national subdivisions are held responsible to the International Society through a Subsecretary, one for each country.

Article Two

PURPOSES

The following are the purposes for which this organization has been set up:

- a) To establish a center of information for phoneticians;
- b) To promote training and research in the field of phonetics in all countries, at all levels in the appropriate branches of the educational system;
- c) To establish contact with kindred scientific organizations and to assist in the establishment of such organizations in countries where there are none.

Article Three

MEMBERSHIP

Sec. 1. Membership of this organization:

- a) Any person interested in the science of phonetics can apply for membership, and on the approval of the Membership Committee (see Article Seven, Sec. 1), become a *regular* member.
- b) International, national, or local organizations pursuing wholly or partly parallel purposes may become *affiliated* members.

The admission of such affiliated members requires the decision of the Board of Directors.

Sec. 2. Any member may at any time withdraw from the Society or may be suspended or expelled by the Board of Directors for any cause prejudicial to the best interests of the Society.

Sec. 3. Membership fees shall be determined by the regular meeting of members at each of the International Congresses of Phonetic Sciences, and on proposition of the Board of Directors. In case such fees or dues, owing to difficulties of currency regulations or for other reasons, cannot be determined on a general scale, the Board of Directors is entitled to settle the question of fees and dues for each country separately. The dues of the affiliated organizations will be determined by the Board of Directors in each case individually.

Article Four

MEETINGS

- Sec. 1. General Meetings to elect the President and the other members of the Board of Directors, for receiving the reports of the directors and for the transaction of other business shall be held at the occasion of, and at the same time as each International Congress of Phonetic Sciences.
- Sec. 2. The regular individual members of the Society form the voting body of its meetings. Every regular individual member is entitled to one vote. All action shall be decided by a majority of the members present.

Article Five

DIRECTORS

- Sec. 1. The business of the Society and the management of its affairs shall be vested in the Board of Directors.
- Sec. 2. Directors shall be elected for a period up to the time of the next General Meeting and may be re-elected by the General Meeting.
- Sec. 3. The Board of Directors shall have control of the property and affairs of the Society and shall fix its policies. They shall have power to hold meetings, organize committees, employ the necessary staff, accept, suspend or expel members as in these by-laws provided, authorize proper expenditures and take all necessary and proper steps to carry out the purposes of this Society and promote its best interests.

Article Six

OFFICERS

- Sec. 1. The officers of the Society shall be the President, one of the Vice-Presidents, the Secretary and the Treasurer. The President shall be elected by the membership directly; the other officers shall be elected by the Board of Directors from among its members. The powers and duties of the officers shall be as follows:

President:

The President shall preside at all regular and special meetings of the Board of Directors and of the Committee, and of the members of the global Society. He shall appoint all committees, with the consent and approval of the Board of Directors; he shall sign all contracts in the name of the Society, enforce these by-laws, cause all books and records to be properly kept and cause all meetings to be called.

Vice-Presidents:

The Vice-Presidents shall perform the duties of President in case of absence of the President. The Board of Directors shall determine which Vice-President shall function as acting President in a particular period.

Secretary:

The Secretary shall keep minutes of all meetings, conduct the correspondence of the Society, and perform all the daily and customary duties of a secretary.

Treasurer:

The Treasurer shall organize and control the accounts of the global and national organizations of the ISPhS, make disbursements only upon proper vouchers signed by the President or a Vice-President, sign checks, make the financial reports of the Society and perform all the customary duties of a treasurer. He shall propose the sum to be contributed every year by each national organization to the Society, expressed as a fractional part of the membership fee paid to the national organizations.

- Sec. 2. No officer shall receive any salary or compensation, except compensation for disbursements made upon instruction of the Board of Directors.

Article Seven

COMMITTEES

- Sec. 1. Membership Committee: The Membership Committee shall review all applications for membership and shall decide about the admission. The Membership Committee is designated by the President on the proposal of the Board of Directors.
- Sec. 2. The Committee of Delegates: The Committee of Delegates is a Consulting body consisting of one representative of each country represented at a Congress and the members of the Society. It discussed all questions concerning the activities of the Society.

Article Eight

AMENDMENTS

These by-laws or any part thereof may be amended, repealed or changed by the vote of a majority of the members at any General Meeting, provided that a copy of the proposed change be mailed to the members at least thirty days before the meeting called to consider that change.

End of By-Laws.

DIE SITZUNG DER INTERNATIONALEN VEREINIGUNG FÜR BIOPHONETIK

In der *Sektion für Biophonetik* waren 16 Vorträge angemeldet, von denen jedoch drei entfallen sind. Eingeleitet wurden die Vorträge durch einen Bericht über die Ergebnisse der von der „Internationalen Vereinigung für Biophonetik“ veranstalteten Enquête über den Schichtenbau des Sprachlautsystems. Organisatorisch wurde die Gründung eines wissenschaftlichen Komitees beschlossen, das aus einer neurologischen und einer biologisch-zoologischen Sektion bestehen soll, die bestimmte aktuelle Probleme der Biophonetik bearbeiten sollen. Die Gründung der an zweiter Stelle genannten Sektion wurde durch die erstmalige Teilnahme von Zoologen an einer biophonetischen Tagung ermöglicht.

Dans la *Section de Biophonétique* on a fait 13 communications qui ont été introduites par un rapport sur les résultats d'une enquête sur la configuration des couches diverses du système du langage. On a décidé la formation d'un comité scientifique qui doit être composé d'une section neurologique et d'une section biológico-zoologique et dont la tâche serait de travailler à certains problèmes de la Biophonétique. La création de la deuxième des deux sections mentionnées a été rendue possible par la présence des zoologues à un colloque de Biophonétique.

RÉUNION DE L'ASSOCIATION INTERNATIONALE POUR LES ÉTUDES PHONOLOGIQUES

La réunion convoquée lundi onze septembre par le comité préparatoire de l'Association internationale pour les études phonologiques a donné son approbation au dit comité sur les relations avec l'Association internationale des Sciences phonétiques.

Comme l'Association phonétique n'a pas encore de structure définitive, on n'a pas effectué l'insertion de l'association phonologique dans la Société phonétique. Mais, on compte le faire à l'avenir. Dans cette perspective, on a coopté dans le comité avec le consentement de la session plénière, le professeur Malmberg, président de l'Association; le comité a été élargi par l'élection du professeur Avanesov (Moscou) et du professeur Hamm (Autriche).

La tâche de ce comité de sept membres (Avanesov, Birnbaum, Hass, Hamm, Malmberg, Martinet, Vachek) sera de préparer une réunion phonologique internationale qui aurait lieu en 1969, à Aix-en Provence.

COMPTE RENDU DE LA SÉANCE DU COLLÉGIUM
DE PHONOLOGIE EXPÉRIMENTALE—PRAGUE
LE 12 SEPTEMBRE 1967

Monsieur le Professeur Malmberg est élu à la Présidence — Monsieur le Professeur Dr. Fry à la Vice Présidence — Monsieur le Docteur Vallancien reste Secrétaire Général et Monsieur Perdoncini Trésorier — Monsieur le Dr. Wustrow Archiviste.

Ont été élus membres: Monsieur le Professeur Luchsinger de Zürich et Madame Sedláčková de Prague.

Monsieur le Professeur Svend Smith suggère l'organisation d'un symposium sur les appareils de synthèse et d'analyse des sons de la parole. Le sujet a été retenu pour une étude ultérieure. Le Congrès se réunira à Paris en Septembre 1968.

COMPTE RENDU DE LA RÉUNION DE I. A. L. S. A.

(International Association of Linguistic Sound Recorded Archives)

Tenu à Prague, le 8 septembre 1867 à 18 h 30

Monsieur le Professeur Zwirner, Président, ouvre cette 4^{ème} réunion des membres de l'Organisation I. A. L. S. A. et remercie les participants pour l'intérêt qu'ils apportent à la discussion des problèmes communs.

Monsieur le Président Zwirner ayant offert sa démission a été solennellement réélu Président à l'unanimité des membres présents.

Monsieur le Président fait un rapport très circonstancié sur l'avancement des publications de la PHONAI, qui avaient pris un léger retard du fait qu'elles avaient demandé la fonte de caractères spéciaux, et dresse la liste des textes et bandes magnétiques de la série allemande actuellement en vente. Monsieur le Dr. Virtaranta, directeur des Archives linguistiques finnoises, présente un très riche exposé sur l'état des enquêtes en Finlande et la méthode à l'aide de laquelle elles ont été réalisées. Les Archives finnoises disposent actuellement de 12.700 heures d'enregistrement, représentant les principales variétés du finnois parlé et de ses dialectes.

Monsieur Richter expose en détail le système I. P. A. G. (dérivé de l'I. P. A.). Le rapport de Monsieur Richter est discuté en séance. Étant donné les qualités du système proposé, de nombreux membres se demandent s'il ne pouvait pas être étendu à toutes les langues représentées au sein de l'association. Monsieur le Président est très sensible à leurs arguments, mais désire auparavant établir une collaboration et une consultation régulière avec la commission de Transcription présidée par Monsieur Isačenko. Le rapport financier est présenté par Monsieur le Dr. Bethge et approuvé à l'unanimité.

R. Gsell, Secrétaire

CLOSING SESSION

D. B. FRY

Mr. President, Ladies and Gentlemen, with your permission I shall not read you a formal report of the meeting of the Permanent Council, but shall follow our usual custom on these occasions and tell you rather more informally the results of the discussions in the Council and also say something about the plans that are being made for the next congress.

Members of the Permanent Council and many members of the Congress have heard with deep regret of the death of two members of the Council. Professor Pierre Fouché was an Honorary Member of the Council: news of his death reached us here in Prague and he will be sadly missed by all those who knew him. It was a great sorrow to us all to hear also of the death, at a comparatively early age, of Professor Gordon Peterson who was an Active Member of the Council and a most gifted research worker in our field. His untimely death means a grievous loss to the phonetic sciences and to many of us personally who were proud to be numbered among his friends.

At the meeting of the Permanent Council, Professor Eberhard Zwirner told us that he thought the time had come for him to retire from active membership. Professor Zwirner has been for very many years a very active Active Member and has helped the cause of the phonetic sciences in many different ways. It was with the very greatest regret that the Council felt obliged to respect Professor Zwirner's wishes in this matter and to accept his resignation. We should like him to know how sincerely grateful we are for all the help he has given in the past. Professor Zwirner has accepted nomination as an Honorary Member of the Council and we are happy to know that we shall still have the benefit of his wisdom and advice.

It is recommended that the two vacancies in the Permanent Council should be filled by the appointment of Professor Kenneth Pike of the University of Michigan and Professor Gerold Ungeheuer of the University of Bonn as Active Members.

I come now to the matter of the Seventh International Congress of Phonetic Sciences, and in this connection the first point discussed by the Council was the date of the next congress. It was decided that a suitable interval would be four years and our next meeting will therefore be held in 1971. At the time of the Council meeting we had not received a firm invitation from any university but I am very happy to be

able to tell you that our friends in Canada have made the suggestion that the congress should be held in Montreal. Professor Rigault and his colleagues have still to discuss the possibility with their university and to seek support for the proposal in Canada but we are most grateful to them for the suggestion and we trust that it will be possible for us to meet in Montreal in 1971.

You would not wish me, I am sure, to conclude these remarks without once more tendering our warmest thanks to our hosts in Prague. To Professor Hála, Professor Romportl and to all their colleagues we want to say that we are most deeply grateful for all their efforts. This has been from every point of view an extremely successful congress, excellently organised, full of interest and highly enjoyable and we do want to thank our friends in Czechoslovakia most sincerely for the part they have played in making this possible.

PARTICIPANTS

- Abramson, Arthur S.*, Professor, Dept. of Linguistics, University of Connecticut, Conn. 06268, USA
- Achmatov, Toktosun K.*, Kirgizskaja SSR, Frunze 24, ul. Frunze 547, USSR
- Ackermann, Ruth*, Dipl. Sprecherzieherin, Universitäts-HNO-Klinik. Halle/S., Leninallee 12, DDR
- Adamson, Vera, B. A.*, Lecturer, Dep. Engl. Lang. & Lit., University Birmingham. Birmingham 15, Great Britain
- Ajrapetjan, V. N.*, Institut russkogo jazyka, Akademia nauk SSSR, Moskva G 19, Volchonka 18/2, USSR
- Akhmanova Olga*, Prof., Moscow G — 69, Ulica Čaikovskogo 16, fl. 15, USSR
- Akin, Johnye*, Professor, Dep. of Speech, University Park, Denver, Colorado 80210, USA
- al-Ani, Salman H.*, Dr. Institute of Islamic Studies, McGill University, Montreal, Quebec, Canada
- Albrow, K.*, Dep. of Linguistics, University College London, Gower Street, London W.C.1, Great Britain
- Altevojt, R.*, Professor Dr., Zoologisches Institut, 44 Münster, Badenstrasse 9, BRD — Frau Dr. Rosamunde Altevojt
- Amengual Gloria*, Dr. Prof., Universidad de la Habana, Cuba (Delegate of the Universidad de la Habana)
- Annan, Brian W. R.*, Lecturer, Dep. of Phonetics, Univ. Leeds, Leeds 2, Yorkshire, Great Britain
- Annuss, Walter*, Dr., 8 München 5, Reichenbachstr. 9, BRD
- Antipova, Ant. M.*, Kandidat filologičeskich nauk, Moskva G 34. Metrostrojevskaja 38, USSR
- Antonelli, Antonio*, Dr., Clinique Oto-rhino-laryngologique de l'Université Arcispedale S. Anna, Ferrara, Italy
- Anttila, Raimo*, assistant professor, Dept. of Classics, Section of Indo-European Studies, Univ. of California, Los Angeles, California 90024, USA
- Arndorfer, Herbert*, Dr., Director of German Instruction, A 1050 Wien, Siebenbrunnengasse 5 a/13, Austria — Frau dr. Edda Arndorfer

Artemov, Vladimir Alekseevič, Prof., Moskva G — 146, 2-aja Frunzenskaja, 10/2, kv. 185, USSR

Augerot, James E., Professor, Instructor, University of Washington, Dept. of Far Eastern and Slavic Languages and Lit., Seattle, Washington 98105, USA

Auteserre, Denis, Collaborateur technique CNRS, Laboratoire de Phonétique, Parc la Provence B 6 13, Marseille 12e, France — Mme France Auteserre, Mlle M. Auteserre

Auty, Robert, Professor, Comparative Slavonic Philology, Brasenose College, Oxford, Great Britain

Avanesov, Ruben I., Professor, Institut ruskogo jazyka AN, Volchonka 18/2, Moskva G-19, USSR (Delegate of the Academy of Sciences USSR)

Avram, Andrei, Dr., Centrul de cercetări fonetice și dialectale al Academiei R. S. România, Bucarest, Calea Victoriei 194, Roumanie

Bachman, Luděk, Dr., CSc., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia

Bagmut, Alla Josifovna, AN Ukrainskoj SSR, Institut jazykovedenija im. L. L. Potebni, Kijev, Kirova 4, USSR

**Baitchura, Uzbek*, Dr., Leningrad, Smolenskaja 1, kv. 15, USSR

Baker, Eleonora, 43 East 78 Str., New York, N. Y. 10021, USA

Baláž, Peter, Filozofická fakulta UK, Bratislava, Gondova 2, Czechoslovakia

Baldán, Alberto, Argine Sinistro 3, Dolo, Venezia, Italy

Baldán, Gino, M. D. Prof., Clinica Otorinolaringoiatrica, Università Padova, Italy

Baldwin, John, Lecturer, Dept. of Phonetics, University College London, London W.C.1, Gower Str., Great Britain

Balkenhol, Heinz, 2 Hamburg 13, Rentzel Str. 17, Alfred-Delp-Haus, BRD

Bamgbose, Ayo, Lecturer, University of Ibadan, Ibadan, Nigeria

Bartoš, Lubomír, Dr., Šumavská 30, Brno, Czechoslovakia

Bartošek, Jaroslav, Dr., Olomouc, FFUP, Křížkovského 10, Czechoslovakia

Baryšnikova, Kira Kipljakovna, Professor, Moskva G — 34, Metrostrojovskaja 38, USSR

**Baštecký, Jaroslav*, Dr., CSc., Psychiatric Research Institute, Prague 8-Bohnice, Czechoslovakia

Bazovská, Elena, Lecturer, Pedagogická fakulta, katedra cudzích jazykov, Bánska Bystrica, Czechoslovakia

Bazytko, Slavomir, Lecturer of Warsaw University, Warszawa, Niecala 6 m 44, Poland

Becker, Ruth, Professor Dr., Ambulatorium für Stimm- und Sprachgestörte des Instituts für Sonderschulwesen der Humboldt-Universität zu Berlin, 104 Berlin, Schumannstr. 20—21, DDR

* not present in person

**von Beckerath, Thea*, Dipl. Psych., Lehrbeauftragte a.d. Universität Bonn, Bad Godesberg, Kurfürstenstr. 5, BRD

Bédard, Albert, Professeur de Phonétique, 78 rue de Sèvres, Paris VIIe, France

Bělič, Jaromír, Professor Dr., Corresp. Member of the Czechoslovak Academy of Sciences, Praha 4-Braník, Ke Krči 22, Czechoslovakia

Beneš, Eduard, Dr., CSc., Ústav jazyků a literatur ČSAV, Praha 1, Valentinská 1, Czechoslovakia

van den Berg, Berend, Professeur à l'Université d'Utrecht, Wevelaan 71, Utrecht, Holland (Delegate of the Utrecht University)

Bergmann, Gunter, Dr., Diplomgermanist, wiss. Oberassistent, Tschairowskistr. 18, 701 Leipzig, DDR

Bergsveinsson, Sveinn, Professor Dr., 112 Berlin-Weissensee, Holzkircher Strasse 1 B, DDR

Bethge, Wolfgang, 4400 Münster, Stadtstr. 39, BRD

Beyer, Ernest, Professeur de l'Université de Strasbourg, 7 rue de Burges, 67 Strasbourg, France — Mme Beyer

de Bhaldráithe, Tomás, Professor, University College, Earlsfort Terrace, Dublin 2, Republic of Ireland

Biedermann, Lieselotte, Prof., 28 rue du Maréchal Foch, Haguenau 67, France

Bierwisch, Manfred, Deutsche Akademie der Wissenschaften, 108 Berlin, Otto-Nuschke-Strasse 22/23, DDR

Bínová, Ludmila, Praha 1, Valentinská 6, Czechoslovakia

Birnbaum, Henrik, Professor, Department of Slavic Languages, University of California, Los Angeles, California 90024, USA — Mrs. Marianna D. Birnbaum

Black, John, Professor, Ohio State University, Department of Speech, 154 North Oval Drive, Columbus, Ohio 43210, USA - Mrs. Helen H. Black [Delegate of (1) Speech Association of America, N. Y. and (2) The Ohio State University]

Blanár, Vincent, Docent Dr., vedúci vedecký pracovník ÚSJ SAV, Bratislava, ul. Červenej armády 75, Czechoslovakia

Blanchard, Irene, Dr., School Dept., Pacific State Hospital, Box 100, Pomona, California 91766, USA

Blochina, Ljudmila Petrovna, kandidat filologičeskich nauk, doc., Moskva B-113, Sokolničeskij val, dom 26, korpus 3, kv. 43, USSR

Blom, Johannis G., Engineer in Chief, Institute of Phonetic Sciences of the University of Amsterdam, Herengracht 338, Amsterdam, The Netherlands

Bock, Bernhard, Dr., Höhenblick 1, 33 Braunschweig, BRD — Frau Ursula Bock

Bogdan, Mihail, Professor, Strada Petre Ispirescu 1, University Cluj, Romania

Böhme, Gerhard, Dozent Dr., Universitäts-HNO-Klinik, 701 Leipzig C 1, Liebigstr. 18a, DDR

* not present in person

Bondarko, L. V., Linguist, Leningrad V 164, Universitetskaja nab. 11, Laboratoria eksp. fonetiki, USSR

Borovičková, Blanka, Dr., CSc., Praha 9, Zelenečská 129/28, Czechoslovakia

Bosák, Ján, Ústav slovenského jazyka SAV, Klemensova 27, Bratislava, Czechoslovakia

Boudreau, Marcel, Professeur, Université Laval, Québec, Canada

Bouillon, Jean Pierre, Professeur à I.N.J.S., 254 St. Jacques, Paris 5ème, France — Mme Bouillon

de Bray, Reginald, Professor, School of Slavonic and East European Studies, University of London, London W.C.1, Great Britain

Břida, Jiří, Praha 6, Lotyšská 10, Czechoslovakia

Brend, Ruth M., Dept. of Linguistics, Michigan State University, Wells Hall, East Lansing, Mich. 48823, USA

Brook, Richard, Prof., 3002 Tuscon Drive, Cedar Falls, Iowa, USA

**Broka, Valija Adolfovna*, Valtaiku 16, Riga 17, Latvian SSR, USSR

Brovčenko, Tamara, Philologist, Assistant Professor, Odessa, Pastera 42, Laboratory of Experimental Phonetics, USSR

Brown, Victoria Farr, 1922 Franklin Ave. E., Seattle, Washington 98102, USA

Bryzgunova, Elena Andrejevna, Moskva, ul. Gercena 6, katedra ruskogo jazyka dlja inostrancev gumanitarnych F-TOB, MGU, USSR

Buchtelová, Růžena, Dr., Praha 6, Senegalská 635, Czechoslovakia

Bulovičová, Viera, Dr., CSc., Praha 3, U Rajske zahrady 1a, Czechoslovakia

Bugarški, Ranko, Department of English, University of Belgrade, Takovska 34, Belgrade, Yugoslavia

Burstynsky, Edward N., Professor, Victoria College, University of Toronto 5, Ontario, Canada

van Buuren, Lucas, Lecturer in Linguistics, Linguistics Department, University College of North Wales, Bangor, Wales, Great Britain

van Buuren, Pramila, Mrs., 11, Victoria Park, Bangor, Wales, Great Britain

Buzássyová, Klára, Karlova Ves, Lištie udolie 11, Bratislava, Czechoslovakia

Bystrzanowska, Teofila, Prof. of Otolaryngology, Warszawa, Marszalkowska 68, m. 25. Poland

**Calearo, Carlo, Dr.*, Clinique Oto-rino-laryngologique de l'Université Arcispedale S. Anna, Ferrara, Italy

Capelle, Guy, Professeur, The University of Michigan CRLLB, 220 East Huron Str., Ann Arbor, Michigan 48108, USA — Mrs. Janine Capelle

Carney, Edward, Lecturer, Dept. of Gen. Linguistics, University of Manchester, Manchester 13, Great Britain (Delegate of the University of Manchester)

Carnochan, J., School of Oriental & African Studies, University of London, London W.C.1, Great Britain — Mrs. P. Carnochan, Michael Carnochan

Carton, Fernand, Professeur Dr., Directeur du Laboratoire de Phonétique, Faculté des Lettres, 54-Nancy, France

van Caspel, Paul P. J., Neptunusstraat 39, Groningen, The Netherlands — Frau van Caspel

Cazes, Daniel, Anthropologue-linguiste, Galveston 48, Mexico 18 D. F., Mexico

Cazes, Madalena, Galveston 48, Mexico 18 D. F., Mexico

Ceplītis, L., Turgeneva 19, Riga, Latvia, USSR

**Cerda, Ramon*, Laboratorio de Fonética, Medinaceli 4, Madrid 14, Spain

Chalupová, Věra, Ing., Praha 6, Vokovická 679, Czechoslovakia

Chapallaz, E. Marguerite, M. A., Flat 31, Tavistock Court, Tavistock Square, London W.C.1, Great Britain

Charbonneau, Gérard W., 983 rue Riel, Quartier Duvernay, Ville de Laval, Qué., Canada

Charbonneau, René, Professeur, Université de Montreal, Casier Postal 6128, Canada

Christov Philipp, Bulgarian Academy of Sciences, Bulgarian Language Institute, Sofia, Bulgaria

Chulliat, Roger, Professeur à l'I.N.J.S., 254 rue St. Jacques, Paris 5e, France — Mme Chulliat

Čistovič, L. A., Dr., Pavlov Institute of Physiology, Leningrad, USSR

Clyne, Michael, Dr., Lecturer, Monash University, Clayton, Vic., Australia

Cohen, Marcel, Directeur d'études à l'Ecole des Hautes Etudes, Professeur honoraire à l'Ecole des langues orientales, 20 rue Joseph Bertrand, 78 Viroflay, Yvelines, France — Mme Cohen

Constant, Guy, Ingénieur, 216 rue des Voies du Bois — Pavillon 4, Colombes/Seine, France — Mme Constant

Crahay, Serge, physicien, 151 Boulevard Lambert, Bruxelles 3, Belgique — Mme Crahay

Daan, Jo, Dr., Nieuwe Hoogstraat 17, Amsterdam C, The Netherlands (Delegate of the Dept. of Dialectology, Onomastics and Folklore of the Royal Netherlands Academy of Sciences and Lettres)

Daem, Frans, Lic. Phil. Germ., Berglaan 68, Heverlee, Belgium

Dahl, Dietrich, Dipl. Physiker, 25 Rostock, Doberanerstr. 137/139, DDR — Dr. T. Dahl

Daneš, František, Doc. Dr., Director of the Institute of the Czech Language, Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia

Delattre, Pierre, Professor, Speech Synthesis Project, Bldg. 446, University of California, Santa Barbara, California, USA

Dillon, Myles, Professor, 10 Burlington Road, Dublin 4, Republic of Ireland

Divine, Suzanne, 5113 Baltan Road, Washington 16, D.C., USA

Dokulil, Miloš, DrSc., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia

Dolphyne, Florence, Dr., University of Ghana, Legon, Ghana

Domański, Rimięiusz, Laryngolog, Oberarzt der HNO-Klinik, Warszawa, Marchlewskiego 65/129, Poland

Dostál, František, Ing., Praha 6, Raisova 5, Czechoslovakia
Dravina-Rüke, Velta Tatjana, Doc., Blidvädersvägen 4j, Lund 9, Sweden
Dukiewicz, Leokadia, Dr., Polska Akademia Nauk, Matejki 49, Poznań, Poland
Dunstan, Elisabeth, Donaldson Gardens, St. Andrews, Gr. Britain
Dvořáčová, Jana, Dr., CSc., Rauchova 24, Bratislava, Czechoslovakia
Džaparidze, Zurab Nikolajevič, kandidat filologičeskich nauk, Tbilisi, Dzerdžinskogo 8, Institut jazykoznanija Akademii Nauk Gruzinskij SSR, USSR
Elekfi, László, Budapest V, Szalay u. 10, Nyelvtudományi Intézet, Hungary
Elert, Claes-Christian, Docent, Director of the Dpt. of Phonetics, Institute of Linguistics, University of Stockholm, Box 23 144, Stockholm 23, Sweden
Endres, Werner, Dr. rer. nat., Privatdozent, 61 Darmstadt, Hölderlinweg 27, BRD
Endrjšová, Růžena, Dr., Praha 1, Malostranské nám. 1, Czechoslovakia
Eras, Heinrich, Dipl. Phil., 108 Berlin, Leipziger Str. 3/4, DDR
Eretescu-Golopentia Sanda, Centrul de certări fonetice și dialectale, Clea Victoriei 194, Bucuresti, Romania — Mr. Constantin Eretescu
Eustace, Sinclair, 11 First Street, London SW 3, Great Britain
Ezawa, Kennosuke, Dr., Wissenschaftlicher Assistent an der Univ. 5 Köln, 1, Marsilstein 3—7, BRD — Frau Barbara Ezawa
Fal'han, Francois, Professeur, 26 Rue de Fougères, 35 Rennes, France (Delegate of the Académie de Rennes)
Fant, Gunnar, Professor, Speech Transmission Laboratory, KTH Stockholm, Stockholm 70, Sweden — Mrs. Fant
Fasske, Helmut, Bautzen, Taucherstr. 36, DDR
**Faure, Georges*, Professeur Dr., 181 rue Breteuil, Marseille 6e, France
Feagin, Louise, 207 Poplar Drive, Falls Church, Virginia, USA
Felicetti, Si, Speech Transmission Laboratory, KTH Stockholm 70, Sweden
Fialová, Draň., Křížkovského 10, Olomouc, Czechoslovakia
Filip, Miroslav, Dr., CSc., Department of Musicology of the Komenský University, Kupeckého 11a, Bratislava, Czechoslovakia
Fintoft, Knut, Cand. real., Docent, Sildrøpeveien 17F, 7056 Moholtan, Trondheim Norway — Mrs. Marit Fintoft
Firbas, Jan, Doc. Dr., CSc., Grohova 63, Brno, Czechoslovakia
Fischer-Jørgensen, Eli, Professor, Kongestien 45, Virum, Denmark
Fjelland, Anders, Lektor, Norges Handelshøyskole, Helleveien 30, Bergen, Norway
Flišet, Albert Lange, University Lecturer, Lars Hilles gate 34 E, Bergen, Norway
Flint, Elwyn Henry, University of Queensland, Department of English, St. Lucia, Brisbane, Queensland, Australia
Florján, J., MUDr., Foniatické odd. KÚNZ, ORL odd. nemocnice, Ústí n. Labem, Czechoslovakia
Fournier, Henry, Professeur de Linguistique et Grammaire comparée à la Faculté des Lettres de Bordeaux, 34 Avenue de Mirmont, 33 Bordeaux — Caudéran, France

Fransson, Frans, Civilingenier, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden — Mrs. Ester Fransson
Fredrich, Ruth-Brigitte, wiss. Assistentin, 402 Halle/Saale, Fritz-Reuter Str. 10, DDR
Fried, Vilém, Prof. Dr., CSc., Praha 1, Betlémská 7, Czechoslovakia
Frinta, Ant., Prof. Dr., Na Šafrance 11, Praha 10, Czechoslovakia
Frøkjær-Jensen, Børge, Assistant Professor, Store Kannikestræde 13, University of Copenhagen, Denmark
Fry, D. B., Professor, Department of Phonetics, University College London, Gower Street, London W.C.1, England (Delegate of the University of London) — Mrs. D. B. Fry
Fujisaki, Hiroya, Dr. Ing., Faculty of Engineering, University of Tokyo, Bunkyo-ku, Tokyo, Japan
**Gajdučik, S. M.*, BSSR, Minsk, Zacharova 21, MGPIIJa, Kafedra fonetiki nemeckogo jazyka, USSR
Gaprindašvili, S. G., Professor, Tbilisi, prosp. I. Čavčavadze 2, kvartira 60, Gruzinskaja SSR, USSR
Garmadi, Salah, Chef de la Section de Linguistique de C.E.R.E.S., 23 rue d'Espagne, Tunis, Tunisie
Garvin, Paul L., Professor, Dpt. of Linguistics, 308 Hayes Hall, State University of New York at Buffalo, Buffalo, N. Y. 14214
Geissner, Hellmut, Dr., 66 Saarbrücken 15, Universität des Saarlandes, BRD — Frau Ursula Geissner
Gendron, Jean-Denis, Professeur, Dép. de Linguistique, Faculté des Lettres, Université Laval, Québec, Canada — Mrs. Gisèle B. Gendron
**Genina, Valentina G.*, kandidat filologičeskich nauk, LEFIPR IMGPIIJa im. M. Toreza, Moskva G 34, Metrostrojevskaja 38, USSR
George, Donald, Professor, University of Southern Mississippi, Hattiesburg, Mississippi, USA (Fulbright Lecturer, Beirut, Lebanon) — Mrs. D. George
Geršič, Slavko, Dr., 5 Köln, Salierring 22/24, BRD
Gimson, Alfred Charles, Professor, Phonetics Department, University College, London W.C.1, Great Britain
Glatz, Elke, Diplomspreecherzieherin, HNO-Klinik der Med. Akademie Erfurt, Phoniatische Abt., 50 Erfurt, Nordhäuserstr. 74, DDR
Godfrey, John, Georgetown University, Washington, D.C. 20007, USA
Goodman, Morris, Professor, 621 Foster St. Evanston, Ill, USA
Goossens, Jan, Professor, Waversebaan 154, Heverlee, Leuven, Belgique
Görner, Herbert, Dipl. Phil., Grosse Fleischergasse 21, 701, Leipzig, DDR
Göschel, Joachim, Dr., 355 Marburg/Lahn, Cappel Str. 136, BRD
Gožin, Georgij M., kandidat filologičeskich nauk, LEF IJaII, MSSR, Moskva G 34 Metrostrojevskaja 38, USSR

Green, Peter Stuart, Lecturer, Language Teaching Centre, University of York, Heslington, York, Great Britain

Gregg, Robert, Professor, Dep. of Classics, University of B. C., Vancouver 8, B. C., Canada — Mrs. M. Gregg

De Grève, Marcel, Professeur, 109 avenue Georges Bergmann, Bruxelles 5, Belgique

Grobbe, L. M., K. N. O. Kliniek, Catharijensingel 101, Utrecht, The Netherlands — Mrs. Grobber

Gsell, René, Professeur, Director of the Phonetic Institute, 2 Rue très cloîtres, Grenoble, France — Mme Gsell

Habovštiak, SAV, Bratislava, Czechoslovakia

Hadj-Salah, Abderrahman, Directeur de l'Institut de Linguistique et de Phonétique. Université d'Alger, 67, Bd. Salah Bouakouir, Alger

Hagen, Sigmund, Department of English, Norges Handelshøgskole, Bergen, Norway — Mrs. Hagen

Hála, Bohuslav, Professeur Dr., Praha 6, Puškinovo nám. 2, Czechoslovakia — Mme A. Hálová

**Halle, Morris*, Professor, Department of Modern Languages and Linguistics, Massachusetts Institute of Technology, Room 20C-132, Cambridge, Massachusetts 02139, USA

Hamm, Josip, Professor Dr., Wien, Liebigstr. 5, Slavisches Seminar der Universität, Österreich

Hammarström, Göran, Professor, Monash University, Dept. of Modern Languages, Clayton, Victoria, Australia (Delegate of the Monash University)

Hamp, Eric P., Professor, University of Chicago, Illinois 60637, USA (Delegate of the University of Chicago) — Mrs. Margot Hamp

Handzel, Leon, Dozent Dr., ul. Siemińskiego 17/14, Wrocław, Poland

Hannaford, Reginald L., Assistant Professor, Bowdoin College, Brunswick, Maine, USA

Harms, L. S., Professor, 210 George Hall, University of Hawaii, Honolulu, Hawaii 96822, USA (Delegate of the Pacific Speech Association) — Mrs. Joan Harms

't Hart, J., Instituut voor Perceptie Onderzoek, Insulindelaan 2, Eindhoven, The Netherlands — Mrs. A. S. E. 't Hart-Roest

Hartert, Christianne, 78 Freiburg, Univ. HNO-Klinik, Kilianstr., BRD

Hass, W., Assistant Professor, Dept. of Psychology, University of Chicago, USA

Haudricourt, André, Directeur de recherches au C.N.R.S., 47 rue d'Assas, Paris 6e, France

Havránek, B., Prof. Dr., Member of the Czechoslovak Academy of Sciences, Praha 6, Na Valech 8, Czechoslovakia

Hedblom, Folke, Doc., Director of the Institute for Dialect and Folklore Research, Uppsala, Tegnérgatan 38A, Sweden — Mrs. Hedblom

Heger, Klaus, Professor Dr., Universität Kiel, Esmarchstrasse 68, D-23 Kiel, BRD — Frau Edith-Marion Heger

Heike, George, Dozent Dr., 355 Marburg, Abteilung Phonetik, Forschungsinstitut für deutsche Sprache, Am Krummbogen 28, Block A, BRD

Hendrich, Josef, Maître Assistant à l'Université Charles, Vlašská 13, Praha 1-Malá Strana, Tchécoslovaquie

Herbst, Lieselotte, Dr., Halle/Saale, Talstrasse 37c, DDR

Higginbottom, Eleanor M., Miss, Department of Linguistics, Lindley Hall, I.U., Bloomington, Indiana 47401, USA (Delegate of the University of Cambridge)

Hill, Archibald Anderson, Professor, University of Texas, Box 8120, Austin, Texas, USA

Hladký, Josef, Dr., Arne Nováka 1, Brno, Czechoslovakia

Hlaváč, S., Ing., Fonetická laboratoř ÚJČ, Praha 1, Valdštejnská 14, Czechoslovakia

Hlaváčková, Zdeňka, Ústav pro jazyk český, Praha 1, Letenská 4, Czechoslovakia

Hlavsá, Zdeněk, Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia

Hlubík, Tomáš, Dr., Devonshire Hall, Cumberland Road, Leeds 6, Great Britain

Holm, Claus, Dr., Universitätsklinik für Hals-, Nasen- und Ohrenkranke, 78 Freiburg, Killianstrasse, Abt. Phonaudiologie, BRD

Holt, Jens, Prof., Nordre Ringgade 70, Arhus N, Denmark

Honikman, Beatrice, Lecturer, Dept. of Phonetics, The University, Leeds 2, Yorks., Great Britain

Honty, Ludvík, Dr., Praha 2, Na Výtoni 9, Czechoslovakia

Horák, Gejza, Dr., Ústav slovenského jazyka SAV, Bratislava, Klemensova 27, Czechoslovakia

Horálek, Karel, Prof., Praha 6-Dejvice, Komornická 7, Czechoslovakia

Horecký, Ján, Dr., Ústav slovenského jazyka SAV, Bratislava, Klemensova 27, Czechoslovakia

Hořejší, Vladimír, Dr., CSc., Praha 2, Benátská 3, Czechoslovakia

Hrnčiar, Ján, Ing., Kabinet fonetiky Filozofickej fakulty KU, Bratislava, Gondova 2, Czechoslovakia

Hrušková, Zdeňka, Dr., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia

Huckleberry, W. Alan, Professor, Ball State University, Muncie, Indiana, USA — Mrs. Elizabeth Huckleberry

Iliescu, Maria, Dr., Roma 60, Bucuresti 63, Romania

Izáková, Kristina, prom. hist., Pedagogická fakulta, Nitra, Czechoslovakia

Izakovič, Vesna, Zagreb, Tomislavov Trg 14, Yugoslavia

Jacobson, Rodolfo, Professor, State University of New York, College at Cortland, Cortland, New York 13 045, USA

Jančák, Pavel, Dr., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia

Janota, Přemysl, Doc., Dr. CSc., Praha 4, Hrusická 4, Czechoslovakia

Janotová, Naděžda, logop., Praha 4, Hrusická 4, Czechoslovakia

Januška, Ivo, Ing., Praha 10-Zahradní Město západ č. 2802, Czechoslovakia

Jazayery, Mohammad Ali, Professor, Dept. of Linguistics, University of Texas, Austin, Texas 78 712, USA

Jedlička, Alois, Prof. Dr., Praha 10, Kišíněvská 16, Czechoslovakia

Jelenová, Helena, Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia

Jones, Glyn Evans, Assistant Lecturer, Department of Welsh, University College, Arts Building, Cathays Park, Cardiff, Wales, Great Britain

Jørgensen, Hans Peter, Assistant Professor, Naesborgvej 70, I Hvidovre, Denmark

Jørgensen, Jytte, Christianberg 19II, Fredericia, Denmark

Josková, Marie, Dr., Universita 17. listopadu, Praha 2, Záhřebská 16, Czechoslovakia

Jürgensen, Bengt, Lecturer, 61 Vingards-Allé, Hellerup, Copenhagen, Denmark

Kahl, Hans-Peter, Dipl. Ing., Institut für Allgemeine Nachrichtentechnik, ZH Darmstadt, 61 Darmstadt, Schlossgartenstrasse 8, BRD

Kaiser, Louise, Prof. Dr., Prinsengracht 793, Amsterdam C, The Netherlands

Kamiš, Adolf, Dr., CSc., Praha 10, Bělocerkevská 18, Czechoslovakia

**Kandinskij, B. S.*, LEFIPR, I MGPIIJa im. M. Torea, Moskva, USSR

Kaschube, Dorothea, Professor, Anthropology Department, University of Colorado, Boulder, Colorado 80 302, USA

Kasparova, Margarita Georgijevna, kandidat filologičeskich nauk, Moskva G 34, Metrostrojevskaja 38, USSR

Kaulhausen, Marie-Hed., Dr., Göttingen, Röderer Str. 9, BRD

Kawashima, Atsuo, a. o. Prof., Sakai, Nakamozu 7-cho 1126, B6-205, Osaka, Japan (Kansai-Universität)

Kelly, John, Lecturer, Dept. of Phonetics and Linguistics, SOAS, Malet Str., London W.C.1., Great Britain

Keramitčievski, Slavčo, Professor, Centre of Audiology and Phoniatriy, Skopje, Dunavska 15, Yugoslavia

Keramitčievska, Ljubitza, Professor, Centre of Audiology and Phoniatriy, Skopje, Dunavska 15, Yugoslavia

Khachatryan, Amalia A., Laboratory of Phonetics, Academy of Sciences, Institute of Language, Erevan 1, Abovian Street 15, USSR

Kimball, Janet G., Ass. Professor, Kent State University, Ohio 44240, USA

Kiml, Josef, Dr., Praha 2, Bělohradská 28, Czechoslovakia

Kinsbourne, Marcel, Dr., University of Oxford, Institute of Experimental Psychology, 1 South Park Road, Oxford, Great Britain — Mrs. Kinsbourne

Kiparski, Paul, Professor, Mass. Institute of Technology, Cambridge, Mass. 02139, USA — Mrs. Kiparski

Kitzing, Karin, Timjanvägen 17, Lund, Sweden

Klajman, Stanislaw, Dr., Gdańsk-Wrzeszcz, ul. Jaškowa Dolina 4/2, Poland

Kloster Jensen, Martin, Prof. Dr., Fonetisk Institutt, Langes Gate 1, Bergen, Norwegen

Knetschke, Edeltraud, Dr., 33 Braunschweig, Steinweg 34, Deutsches Spracharchiv, Aussenstelle Braunschweig, BRD

Koch, Walter A., Dr., 4403 Hiltrup bei Münster, Lange Str. 62, BRD

Kohler, Klaus, Institut für Phonetik, Koblenzer Str. 98a, 53 Bonn, BRD

Kohout, Vladimír, Praha 6, Jenerálka 54, Czechoslovakia

Kolmer, Felix, Dr., VÚZORT, Plzeňská 66, Praha 5, Czechoslovakia

Kolodejová, Katarína, logop., OÚNZ Trstená, Czechoslovakia

Konečná, Dana, CSc., Praha 6, Na dlouhém lánu 3, Czechoslovakia

Konopczyński, Gabrielle, Rue Violet „Au Vert Galant“ Bat. A., 25 Besançon, France

Kořenský, J., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia

Koschmieder, Erwin, Professor Dr., 8011 Vaterstetten, Beethovenstr. 42, BRD — Frau Koschmieder

Kossel, Ingrid, dipl. Sprecherzieherin, 40 Halle/Saale, Ludwig Wucherer Str. 69, DDR

Král, Abel, Dr., Gondova 2, Bratislava, Czechoslovakia

Králik, Stanislav, Dr., CSc., Praha 6, Rooseveltova 15, Czechoslovakia

Krámský, Jiří, Dr., Praha 4, Podolská 86, Czechoslovakia

Krech, Eva-Maria, Dr. phil., Institut für Sprechkunde und Phonetische Sammlung der Martin-Luther-Universität, 402 Halle (Saale), Klement-Gottwald-Str. 3, DDR

Kress, Bruno, Professor Dr. phil. habil., 22 Greifswald, Schillstr. 14—15, DDR

Kriaková, Elena, Pedagogická fakulta v Banské Bystrici, Tajovského 8, Czechoslovakia

Krútel, Jozef, Ing., ČSAV, Fyzikálny ústav SAV, Bratislava-Patrónka, Dúbravská cesta, Czechoslovakia

Kubicka, Wanda, Poznań, ul. Swoboda 44, Poland

Kuchař, Jaroslav, Dr. CSc., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia

Kunešová, V., Praha 1, Smetanovo nám. 20, Czechoslovakia

Kuraishi, Goro, Professor, Musashini-Shi, Tokyo, Japan, Kichijoji-Kitamachi 3-5-23

Kurka, Eduard, Docent Dr., Institut für Sprechkunde, Klement-Gottwald-Str. 3., 402 Halle/S., DDR

Laczkowska, Maria, Doc. Dr. med., Poznań, ul. Mickiewicza 22, Poland

Lancia, Roland, Prof., 39 rue Général Ferrié, 38 Grenoble, France — Mme Lancia

Landercy, Albert, Assistant, 17 rue N. Vandercamme, Mignault, Hainaut, Belgique

Lane, Harlan L., Prof., 25 rue Broca, Paris 5e, France

Langová, Jiřina, Dr., Foniatrické odd. FN 2, Praha 2, Žitná 24, Czechoslovakia

Laver, John, Department of Phonetics and Linguistics, University of Edinburgh, Great Britain

Lawrence, Teleté Z., Professor, 3 860 South Hills Circle, Fort Worth, Texas 76109, USA

Lebourhis, Lucien, Université Laval, Dép. de Linguistique, Faculté des Lettres, Québec 10e, Canada

- Lebrun, Y.*, Prof. Dr., 3 rue de l'Eider, Bruxelles 17, Belgique (Delegate of the Unité de Neurolinguistique de l'Université de Bruxelles)
- Ledvinka, František*, RNDr, Praha 6, nám. Svobody 1, Czechoslovakia
- Lee, W. R.*, Dr., 16 Alexandra Gardens, Hounslow, Middlesex, Great Britain
- Lehiste, Ilse*, Professor Dr., Department of Linguistics, Ohio State University, Columbus, Ohio 43210, USA
- Lejska, Ladislav*, Dr., Foniatrické odd. OÚNZ, nemocnice, Kyjov na Moravě, Czechoslovakia
- Lewis, Winslor J.*, Universitátslektor, British Institute, Oslo, Blindern, Norway
- Lieberman, A. S.*, Moskva G 34, Metrostrojevskaja 38, USSR
- **Liiv, Georg*, Experimental Phonetics Laboratory, Institute of Language and Literature, Academy of Sciences, 10 Lenin Avenue, Tallinn, Estonian SSR, USSR
- Lindner, Gernhart*, Prof. Dr., Institut für Phonetik und Kommunikationsforschung, 108 Berlin, Unter den Linden 9—11, DDR
- Lindqvist, Jan*, Civ. ing., Dept. of Speech Communication, Speech Transmission Laboratory, Royal Institute of Technology (KTH), Stockholm 70, Sweden — Mrs. K. Lindqvist
- Liška, Jozef*, Dr., JÚLŠSAV Košice, Czechoslovakia
- Lisker, Leigh*, Professor, Bennet Hall, 342nd Walnut Str., University of Pennsylvania, Philadelphia, USA — Mrs. S. Lisker, Miss L. Lisker, Miss C. Lisker
- Lišková, Olga*, logop., Košice, ul. Dukelských hrdinův, Czechoslovakia
- Lotzmann, Geert*, Dr. phil., Dipl. Sprecherzieher, Lektor an der Universität Heidelberg; 6900 Heidelberg 1, Zur Forstquelle, BRD — Frau Annette Lotzmann
- Lüdtke, Helmut*, Professor Dr., 78 Freiburg i. B., Lorettost. 41, BRD — Frau Sigrid Lüdtke
- Ludvíková, Marie*, Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia
- Luščichina, I. M.*, Laboratorija Inženernoj Psihologii LGU, Leningrad, USSR
- Lutterer, Ivan*, Dr., CSc., Praha 4, Hrusická 2510, Czechoslovakia
- MacCarthy, Peter*, Professor, Dept. of Phonetics, The University of Leeds, Leeds 2, Great Britain
- Mach, Vladimír*, Praha 10-Zahradní Město, západ, blok 1, 2868, Czechoslovakia
- Machmudov, Achmad*, Docent, Gos. Ped. Institut im. Lenina 8, Andižan, Uzbekskaia SSR, USSR
- Mahnken, Irmgard*, Professor Dr., Nikolausberger Weg 31, 34 Göttingen, BRD
- Mahnken, Ortrun*, 46 Dortmund, Hörde Hilgenstrockstr. 8, BRD
- Maiwald, Dietrich*, Dr., IBM Research Laboratory (Speech Processing), Säumerstr. 4, 8803 Rüschlikon-ZH, Switzerland
- Maláč, Vlastislav*, Ing., Praha 3, K lučinám 5, Czechoslovakia
- Malmberg, Bertil*, Professor, Institute of Phonetics, Kävlingevägen 20, Lund, Sweden
- Mańczak, Witold*, Docent, Lobzowska 9, Kraków, Poland
- Maneca, Constantin*, Professor, Piata Kogalniceanu nr. 1, sc. A, ap. 27, Bucarest 1, Romania
- Mançold, Max*, Professor Dr., Passagestr. 7, 66 Saarbrücken 3, BRD
- Mardarescu, Maria*, Calea Victoriei 194, Bucuresti 22, Romania
- Marquardt, William F.*, Professor, University of Illinois, Education Bld, Room 284 F, Urbana, Ill. 61801, USA
- Marshall, John C.*, Medical Research Council, Speech and Communication Research Unit, Univ. of Edinburgh, 31, Buccleuch Place, Edinburgh, Great Britain
- Martens, Peter*, Dr., Universität Hamburg, 2000 Hamburg 13, Von-Melle-Park 6, BRD
- Martinet, André*, Professeur, 3 Place de la Gare, 92 Sceaux, France — Mme Martinet
- Massa, Alessandra*, Dr., Domegliara, Verona, Italy
- Matejka, Ladislav*, Professor, University of Michigan, 2003 Day Street, Michigan 48104, USA
- Mathiot, Madeleine*, 449 Skyeway Road, Los Angeles, California 90049, USA
- Da Matta Machado, Mirian Therezinha*, 9 bd. Jourdan, Pavillon du Liban, Cité Universitaire, Paris 14e, France
- Mattoso-Camara, Joaquim*, Professor, Rua Artur Araripe 110, apt. 205, Gavea, Rio de Janeiro, Brazil
- Mazzarella, Silvester*, Lecturer, Hackington House, St. Steven's Hill, Canterbury, Kent, Great Britain
- McConnell, Freeman*, Professor, Division of Audiology and Speech Pathology, Vanderbilt University School of Medicine, Nashville, Tennessee 37203, USA
- McDavid, Raven*, Professor, University of Chicago, Chicago, Illinois 60637, USA
- Mehnert, Dieter*, Ing., 117 Berlin, Zur Nachtheide 14, DDR
- Meier, Georg F.*, Professor Dr., Direktor des Instituts für Phonetik und Kommunikationswissenschaft, 110 Berlin — Pankow, Kavalierstr. 19, DDR
- Meinecke, Fritz*, Akad. Oberrat, 6 Frankfurt, A. M. Mühlacherstr. 13, BRD
- Meinhold, Gottfried*, Dr. 69 Jena, Käthe-Kollwitz-Str. 8, DDR
- Meinsma, Gerrit L.*, Herengracht 338, Amsterdam, The Netherlands
- Mejstřík, Vladimír*, Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia
- Melika, Georgij*, Užgorod, USSR
- **Melnikov, Gennady*, Ing., Moskva D — 98, Žukovskij projezd, dom 20, K 3, kv. 26, USSR
- Merhaut, Josef*, Professor Dr. Ing., Praha 10, Solidarita F VI-8, Czechoslovakia
- Merlingen, Weriand*, Dozent Dr., 1170 Wien, Rudolf-Bärenhartg. 15, Österreich
- Michajlovič, Ljubomir*, Professor, University of Beograd, 14, Topličin venac, Beograd, Yugoslavia
- Mikolauskaite, Elzbieta*, Gosudarstvennyj pedagogičeskij institut, Vilnius, Studentu 39, USSR

Mikus, Radivoj, Professeur Dr., Resljeva 4/II., Ljubljana, Yougoslavie — Mme Aleksa Kotnik-Mikus
Mimrová, Milada, Praha 4, Budějovická 800, Czechoslovakia
Minow, Martin, Studenthemmet Tempus, Rum 1316, Beckombergvägen 3, Bromma, Sweden
Mol, H., Prof. Dr. Ir., Institute of Phonetic Sciences of the University of Amsterdam, Herengracht 338, Amsterdam, The Netherlands — Mrs. C. Mol
Morávek, Milan, Dr. Praha-Bubeneč, Wintrova 5, Czechoslovakia
Müller, Horst, Institut für Sprechkunde, 402 Halle, Klement-Gottwald/Straße 3, DDR
Nábělek, Igor, Doc. Dr., Fyzikální ústav SAV, Bratislava-Patrónka, Dúbravská cesta, Czechoslovakia
Nedvědná, Milada, Dr., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia
Neiland, Mirdza A., CSc., Riga 39, 10, Struktora, flat 4, USSR
Neustupný, Jiří V., Professor, Orientální ústav ČSAV, Lázeňská 4, Praha 1, Czechoslovakia
Newell, Alan, el. engineer, Standard Telecommunication Laboratories, London Road, Harlow, Essex, Great Britain
Nickel, G., Professor Dr., 7 Stuttgart 1, Goethestrasse 2, BRD
Novák, Alexej, Dr., Praha 2, Žitná 24, Czechoslovakia
Novák, Ludovít, Prof. Dr., Lubochňa, Czechoslovakia
Novák, Pavel, Docent Dr., Praha 4-Spořilov II, Hlavní 2729, Czechoslovakia
Novotná, Jiřina, Dr., CSc., Praha 10, U vršovického hřbitova 572, Czechoslovakia
Nowicki, Jan, Dr. med., Bukietowa ul. 12, m. 6, Warszawa-Mokotów, Poland
Nurmachanova, Aziza Nurmachanovna, Dr., Moskva G 34, Metrostrojevskaja 38, USSR
Öhman, Sven E. G., Prof. Stockholm, Flintbacken 6, Sweden — Mrs. B. Öhman
Ohnesorg, Karel, Professor Dr., Brno 2, Šumavská 36, Czechoslovakia (Delegate of the J. E. Purkyně University)
Oliverius, Zdeněk, Dr., CSc., Lecturer, Praha 10-Zahradní Město, Práčská ul. 2672, Czechoslovakia
Ondráček, Jaroslav, Nové Město na Moravě, Pionýrů 440, Czechoslovakia
Ondráčková, Jana, Dr. CSc., Praha 10-Strašnice, Nad Primaskou 15, Czechoslovakia
Onishi, Masao, Professor, 13 Daita-2, Setagayaku Tokyo, Japan
Oravec, Ján, Dr., Bratislava, Klemensova 27, Czechoslovakia
Pačesová, Jaroslava, Doc., Dr. CSc., Fanderlíková 72, Brno 16, Czechoslovakia
Pahn, Johannes, Dr., Erfurt, Med. Akademie, HNO-Klinik, Nordhauser Str. 74, BRD — Frau Pahn
Palek, Bohumil, Dr., Praha 3, Rečkova 5, Czechoslovakia
Palková, Zdena, Dr., Praha 3, Rečkova 5, Czechoslovakia
Parlangèli, Oronzo, Professor, University of Bari, Via G. Brunetti 51, Novoli (Lecce),

Italy Delegate of the (1) Università di Bari, (2) Gruppo di Ricerce per la Dialettologia Italiana (Consiglio Naz. delle Ricerce)
Pascher, Wolfgang, Dr., Oberarzt, 2000 Hamburg 20, Schott-Müllerstr. 1/IV., BRD
Paška, Martin, Pedagogická fakulta, Martin, Czechoslovakia
Paufošima-Švanberg, Rozalija Francevna, Institut ruskogo jazyka AN SSSR, Moskva G 19, Volehonka 18/2, USSR
Pauliny, Eugen, Professor Dr., Bratislava, Poštová 4, Czechoslovakia (Delegate of the Komenský University, Bratislava)
**Paulsen, Karsten R.*, Privatdozent und Oberarzt, Hals-Nasen-Ohren-Klinik der Universität Kiel, 23 Kiel, Hospitalstr. 20, BRD
Paulus, Erwin, Institut für Datenverarbeitung, Techn. Hochschule München, Arcisstr. 21, BRD
Pazderka, Zdeněk, Dr., primář ORL odd. OÚNZ nemocnice, Příbram, Czechoslovakia
Pech, Karel, Dr., Praha 6-Vokovice, Na červeném vrchu 674/12, Czechoslovakia
Peciar, Štefan, Dr., CSc., Bratislava, Jazykovedný ústav E. Štúra SAV, Klemensova 27
Peco, Asim, Docent Dr., Studentski trg 3, Beograd, Yugoslavia
Pée, Willem, Professor, Vuursteendreef 3, Bosvoorde, Bruxelles 17, Belgique (Delegate of the Université de Ghent) — Mme Pée
Penzl, Herbert, Professor, Department of German, 5319 Dwinelle Hall, University of California, Berkeley, Calif. 94720, USA — Mrs. Penzl
Peter, Harry Winfield, Lecturer in Linguistics, La Moyne College, Syracuse, New York 13214, USA
Petitclerc, Grace, 1550 Wilder Ave., Honolulu, Hawai 96822, USA
Petřík, Miroslav, Ing., Praha 2, Žitná 24, Czechoslovakia
Pheby, John, Linguist, Deutsche Akademie der Wissenschaften, Leipzigerstr. 3—4, 108 Berlin, DDR
Pignataro, Oresto, M. D., Ass., Clinique Universitaire ORL, Università di Ferrara, Italy, — Mme O. M. Pignataro
Pike, Kenneth L., Prof., Dept. of Linguistics, 214 Gunn Bldg., Univ. of Michigan, Ann Arbor, Mich. 48104, USA — Mrs. Pike
Poch, Robert, M. D., Department of Radiology, Faculty of Medicine, Šrobárova 50, Praha 10, Czechoslovakia
Podolcová Gabriela, odb. as. FFUK, Bratislava, Kupeckého 9a, Czechoslovakia
Poštolková, B., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia
Potapova, Rodmonga, Prof., Moskva, A-445, Leningradskoje šosse, dom 112 1, korpus 3, kv. N 609, USSR
Pražič, Mihajlo, Prof. Dr., Audiological Department of the Ear, Nose and Throat Clinic, Medical Faculty, Zagreb, Yugoslavia
Pulgram, Ernst, Professor, University of Michigan, Ann Arbor, Mich. 48104, USA
Raabe, Margret, Dipl. Sprecherz., Halle/S., Schillerstr. 23, DDR
Radová, Magdalena, Fonetická laboratoř ÚJČ, Praha 1, Valdštejnská 14, Czechoslovakia

von Raffler Engel, Walburga, Professor Dr., Vanderbilt Univ., Nashville, Tennessee 37203, USA (Delegate of the Vanderbilt University)
Rainov, Vasil, Dr., Krivolak 3, Sofia Bulgaria
Reinhold, Ernest, Dr., Head, Dept. Germanic Languages and General Linguistics, University of Alberta, Canada (Delegate of the Canadian Linguistic Association)
Rensch, Karl H., Dr., Wiss. Assistent am Institut der Universität Köln, 5 Köln, Salierring 22/24, BRD
Revtova, Ljubov Dmitrijevna, Moskva G 34, Metrostrojevskaja 38, USSR
Richter, Helmut, Diplom-Psychologe, 5 Köln, Salierring 22—24, BRD — Frau Brigitte Richter
de Ridder, Peter, Publisher, Frederik van Eedenlaan 61, Leiden, The Netherlands
Rigault, André, Directeur, Département de Linguistique, Université McGill, Montréal 2, Canada
Ringgaard, Kristian, Professor, dr. phil., University of Aarhus, Aarhus, Denmark — Mrs. Ringgaard
Risberg, Arne, Civ. ing., Speech Transmission Laboratory, Dept. of Speech Communication, Royal Institute of Technology (KTH), Stockholm 70, Sweden
Rischel, Jørgen, Assist. Prof., Skyttebjerg 123, Nærum, Denmark — Mrs. Anna-Grethe Rischel
Roceric-Alexandrescu, Alexandra, Scientific Researcher, Str. Muntü Tatra, ur. 3, Rn. 30 Decembrie, Bucuresti, Romania
Robson, W. H., Univ. Lecturer, Bjerkebakken 68 A, Oslo 7, Norway
Rodon, Eulalia, Prof., Instituto „Maria de Austria“, Antonio de Layva 64, Madrid 19, Spain
Romportl, Milan, Professor Dr., Praha 6, Havlovská 40, Czechoslovakia (Delegate of the Charles University, Prague) — Frau L. Romportlová
Romportl, Simeon, Dr., Faculty of Pedagogics, J. E. Purkyně University, Brno, Poříčí 7, Czechoslovakia
Rossi, Mario, Dr., 19 Bd. Joseph Fabre, Marseille 12e, France, (Delegate of the Faculté des Lettres d'Aix en Provence) — Mme Marguerite Rossi
Rossi, Mario, Clinica Otorinolaringoiatrica, Università di Padova, Italia
Roudný, Miroslav, Dr., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia
Runneburger, Henri, Ecole Louis Gros, 84 Avignon, France
Sabol, Ján, Dr., Prešov, Pavlovičovo nám. 42, Czechoslovakia
Šabršula, Jan, Prof. Dr., Praha 1, Jáchymova 4, Czechoslovakia
Sagen, Trygve, University Lecturer, Bergsalleen 8, Oslo 8, Norway
Saltarelli, Mario, Professor, Div. of Modern Languages, Cornell University, Ithaca, New York 14850, USA
Sasaki, Miyoko, Ass. Prof., Tokyo, Shibuya-ku 34, 2, Nishihara, Japan

Šaumjan, K. S., Dr., Institut russkogo jazyka AN, Moskva, USSR
Schädlich, Hans-Joachim, Dr., 108 Berlin, Deutsche Akademie der Wissenschaften, Leipziger Str. 3/4, DDR — Frau Krista Maria Schädlich
Schindler, Oskar, Professor, Turin, C. Turati 82, Italy
Schneider, Alfred, Professeur, 85 rue d'Havré, Mons, Belgique
Schneider-Kobersky, Eva, 85 rue d'Havré, Mons, Belgique
Schnorrenberg, Josef E., cand. phil., 5 Köln, Salierring 22/24, Institut für Phonetik, BRD
Schouten, J., Groningen Univ., Kerkduinweg 14, Noordlaren, Holland — Frau J. C. M. Schouten-Elzinga
Schubiger, Maria, Dr., Lecturer, 81 Hochstrasse, 8044 Zürich, Schweiz
Schultink, H., Prof. Dr., Instituut A. W. de Groot voor Algemene Taalwetenschap, Wilhelminapark 12, Utrecht, The Netherlands
Schulzová, Olga, Bratislava, Februárového víť. 13/VI, Czechoslovakia
Schwanzer, Viliam, Docent Dr., Bratislava, Šoltésovej 10, Czechoslovakia
Schwanzerová, Jolana, Bratislava, Šoltésovej 10, Czechoslovakia
Schwenk, Sigrid, Dr., 355 Marburg/Lahn, Forschungsinstitut für deutsche Sprache, Deutscher Sprachatlas, Kaffweg 3, BRD
Scully, Celia, Dept. of Phonetics, University of Leeds, Leeds 2, Yorkshire, Great Britain
Sebestyén, Árpád, Dozent, Universität L. Kossuth, Debrecen, Ungarn
Sedláček, Karel, Prof. Dr., Praha 2, Karlovo nám. 18, Czechoslovakia
Sedláčková, Eva, Dozent Dr., Praha 2, Karlovo nám. 18, Czechoslovakia
Seeman, Miloslav, Professor Dr., Praha 2, Nové Město, nábř. B. Engelse 66, Czechoslovakia
Seiffert, Leslie, Lecturer, Dept. of German, University of Birmingham, Birmingham 15, Great Britain
Sgall, Petr, Docent Dr., Praha 4, Olbrachtova 1049, Czechoslovakia
Shen, Yao, Professor Dr., University of Hawaii, Honolulu, Hawaii 96822, USA
Short, David, B. A., 23 Cheltenham Ave, Liverpool 17, Great Britain
Silva, Georgette, University of California, Los Angeles, California 90024, USA
Simon, Péla, Dr., 33 Bd. d'Anvers, 67 Strasbourg, France
Singh, Sadanand, Howard University, Washington D.C., USA
Sivertsen, Eva, Professor Dr., Møllebakken 36, Trondheim, Norway
Skalička, Vladimír, Professor Dr., Praha 2, Dřevná 6, Czechoslovakia
Skaličková, Alena, Doc. Dr., CSc., Praha 2, Dřevná 6, Czechoslovakia
Skoumal, Jan, Praha 10-Zahr. Město, Kapradová 2649, Czechoslovakia
Slama-Cazacu, Tatiana, Rahovei 108, Bucuresti, Romania
Slis, I. H., Instituut voor Perceptie Onderzoek, Insulindelaan 2, Eindhoven, The Netherlands — Mrs. J. W. Slis-Smienk
Smith, Svend, Professor Dr., Phonetisches Institut, Alsterglaciis 3, Hamburg, BRD
Söderberg, Carl-Gustav, Professor, Erikstorpsgatan 10 B, Malmö V, Sweden

Sonninen, Aatto, M. D., Gummeruksenk. 3. B. 24. Jyväskylä, Finland
Sosna, Fr., Dr., Praha 3, Písecká 9, Czechoslovakia
Sovák, Miloš, Professor Dr. med. Dr. phil., Praha 1, Kaprova 14, Czechoslovakia
Sovijärvi, Antti, Professor Dr., Institute of Phonetics, Vironkatu 1 B 15, Helsinki 17, Finland
Spang-Thomsen, Borge, Søndervangen 86, 3460 Birkerød, Denmark
Spencer, Margaret, Professor Dr., Asian and Pacific Languages Dept., University of Hawaii, Honolulu, Hawaii 96822, USA
Sperlbaum, Margret, Dr., 33 Braunschweig, Steinweg 34, Deutsches Spracharchiv, Aussenstelle Braunschweig, BRD
Stageberg, Norman C., Prof., Dept. of English, University of Northern Iowa, Cedar Falls, Iowa, USA
Stelzig, H., Doz. Dr. habil., Univ. Greifswald, Abt. Sprechwissenschaft, Löns-Str. 20/I, DDR
Stewart, John M., Dr., Institute of African Studies, University of Ghana, Legon, Ghana
Stich, Alexandr, Dr., CSc., Ústav pro jazyk český, Praha 1, Letenská 4, Czechoslovakia
Stock, Eberhard, Dr., Institut für Sprechkunde, Kl. Gottwald Str. 3, 402 Halle, DDR
Stojanović, Časlav, Institut za strane jezike, Gospodar Jovanova 35, Beograd, Yugoslavia
Stojkov, Stojko, Professor Dr., Sofia 6, F. Kanitz 38, Bulgaria
Štolc, Jozef, Professor Dr., Bratislava, Trnavská 7, Czechoslovakia
Storms, Godfried, Professor Dr., Prinses Margrietlaan 4, Berg-en-dal (bij Nijmegen), The Netherlands (Delegate of the Katholieke Universiteit, Nijmegen) — Mrs. M. J. M. Storms-Wilmink
Šupáček, Ivan, Dr., Praha 1, Kaprova 14, Czechoslovakia
Șuteu, Valeriu, Centrul de cercetări fonetice și dialectale, Bucarest, Calea Victoriei 194, Romania
Suttner, Jutta, Sprecherzieherin Lektorin, Halle/Saale 402, Am breiten Pfuhl 42, DDR
Švarný, Oldřich, Dr., CSc., Orientální ústav ČSAV, Praha 1, Lázeňská 4, Czechoslovakia
Švestková, Ludmila, Dr., Praha 6, Bělohorská 141, Czechoslovakia
Sychra, Ant., Prof. Dr., Praha 2, Plavecká 8, Czechoslovakia
Takefuta, Yukio, The Ohio State University, Dept. of Speech, 154 North Oval Drive, Columbus, Ohio 43210, USA — Mrs. Katsuko Takefuta
Taranovski, Kiril, Professor, Harvard University, Boyston Hall 301, Cambridge, Mass. 02138, USA — Mrs. Vera Taranovski
Tataru, Ana, Cluj, Str. Porumbeilor 34, Romania
Tatham, Marcel A. A., University of Essex, Colchester, Essex, Great Britain
Teatini, Gianpietro, Dr., Otolaryngological Clinic of the University Arcispedale S. Anna, Ferrara, Italy — Mrs. X. Teatini

Tembrock, Günther, Professor Dr., 104 Berlin N 4, Zoologisches Institut der Humboldt-Univ., Invalidenstr. 43, DDR
Tervoort, Bernard Th., Professor Dr., Bussum, Prins Hendriklaan 7, The Netherlands
Těšitelová, M., Dr., CSc., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia
Theiner, Christa, Dr., Wiss. Mitarbeiterin, 104 Berlin, Institut für Sonderschulwesen, Albrechtstr. 22, DDR
Thomas, Ceinwen H., Dr., Welsh Language Research Unit, Department of Welsh, University of Wales, Cardiff University College, Cathays Park, Cardiff, Wales, Great Britain
Thorsen, Ofuf M., Assistant Professor, Institute of Phonetics, University of Copenhagen, St. Kannikestraede 13, Copenhagen, K. Denmark
Tichý, Jiří, Doc. Dr., Praha 5, Vančurova 23, Czechoslovakia
Tilkov, Dimitar, 55 rue Iskar, Sofia, Bulgaria
Tillmann, H. G., Dr., D 8000 München 23, Biedersteiner Str. 59, BRD
Tolk, J., Drs., Audiological Department of the University of Nijmegen, St. Annastraat 315, Nijmegen, The Netherlands
Tománek, Rostislav, MUDr., ORL Laboratory of ČSAV, Praha 2, U nemocnice 2, Czechoslovakia
Tomanović Vaso, Prof. Dr., Leninova 16-a, Skopje, Yugoslavia
Toporišič, Jože, Dr., Filozofska fakulteta, Aškerčeva 12, Ljubljana, Yugoslavia
Torsujeva, Irina G., Moskva G 34, Metrostrojevskaja 38, USSR
Trim, John, Lecturer, Selwyn College, Cambridge, Great Britain
Trnka, Bohumil, Prof. Dr., Praha 1, U železných lávkách 16, Czechoslovakia
Trojan, Felix, Prof. Dr., Wassergasse 15, A 1030 Wien III, Österreich
Trojanová, Jar., Dr., Radlas 16, Brno, Czechoslovakia
Tronczyńska, Jadwiga, Dr., Łódź, Kopcińskiego 22, Klinika Laryngologiczna, Poland (Delegate of the Polish Otolaryngological Society)
Trost, Pavel, Prof. Dr., Praha 6, Na Míčáncích 20, Czechoslovakia
**Truby, Henry M.*, Dr., University of Miami, Coral Gables, U. of M. Boz 8085, Florida 33124, USA
Tscheschner, Walter, Dr. Ing., Pillnitzer Str. 24, 801 Dresden, DDR — Frau Ruth Tscheschner
**Turkbenbajev, N.*, LEF IJa AN KazSSR, Alma Ata, USSR
Ulbrich, Horst, Dr., 1193 Berlin-Treptow, Köpenicker Landstr. 54, DDR
Ungeheuer, Gerold, Prof. Dr., Institut für Phonetik und Kommunikationsforschung der Universität Bonn, 53 Bonn, Adenauerallee 98a, BRD
Urbelene J., LTSR, Vilnius, Verkiu, 20—51, USSR
Vachek, Josef, Professor Dr., Praha 3, nám. Jiřího z Poděbrad 18, Czechoslovakia
Valdman, Albert, 310 Lindley Hall, Indiana University, Bloomington, Indiana 47401, USA — Mrs. Hilde Valdman
Valentin, Paul, Dr., 27 rue Henri-Savignac, 92 Moudon-Bellevue, France, — Mme Valentin

- Vallancien, Bernard*, Dr., 16 rue Spontini, Paris 16e, France
- Vanek, Anthony L.*, Department of Linguistics, 309 Davenport Hall, University of Illinois, Urbana, Illinois 61801, USA
- Vaniček, František*, Ing., Praha 6, Angolská 610, Czechoslovakia
- Fanvik, Arne*, Lecturer, Sofienbergstr. 63B, Oslo 5, Norway
- Vařecha, Vladimír*, Dr., CSc., Univ. 17. listopadu, Praha 1, Czechoslovakia
- Vašák, P.*, Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia
- Vašek, Ant.*, Doc. Dr., CSc., Brno, Úvoz 59, Czechoslovakia
- Vasilău, Emanuel*, Centrul de cercetări fonetice și dialectale, Calea Victoriei 194, București, Romania
- Vasiljev, V. A.*, Professor, LEFIPR I MGPIIJa, Moskva G 34, Metrostrojevskaja 38, USSR
- Vazquez Cuesta, Pilar*, Menéndez Pelayo 63, 6° B — Madrid 9, Spain
- Vecozol, Marta Pavlovna*, ul. Suvorova 109, kv. 5, Riga 1, Latvijskaja SSR, USSR
- Vende, Kullo*, Experimental Phonetics Laboratory, Institute of Language and Literature, Academy of Sciences, 10 Lenin Avenue, Tallinn, Estonian SSR, USSR
- Verbickaja, L. A.*, Leningrad, Universitetskaja nab. 11, USSR
- Verburg, Pieter A.*, Professor Dr., Hagepreekkamp 4, Groningen, The Netherlands — Mrs. Verburg
- Viereck, Wolfgang*, Dr., Seminar für Englische Sprache und Kultur der Universität Hamburg, 2 Hamburg 13, Von-Melle-Park 6, BRD
- Vinař, Oldřich*, Dr., CSc., Výzkumný ústav psychiatrický, Praha 8, Bohnice, Czechoslovakia
- Virtaranta, Pertti*, Prof. Dr., Risto Rytin tie 28A, Helsinki, Finland
- Višnevskaja, Lidija Jakovlevna*, Moskva G 34, Metrostrojevskaja 38, USSR
- Vrbová, Libuše*, Fonetická laboratoř ÚJČ, Praha 1, Valdštejnská 14, Czechoslovakia
- De Vriendt, Séraphin*, Chargé de Cours, 14 rue Gulden Bodem, Bruxelles 8, Belgique
- Vrtička, Karel*, Dr., CSc., Praha 2, Žitná 24, Czechoslovakia
- Wajskop, Mazmilien*, Directeur, Institut de Phonétique, 50 avenue F. D. Roosevelt, Bruxelles 5, Belgique (Delegate of the Université libre de Bruxelles)
- Webster, John C.*, US Navy Electronic Laboratory, Code 3170, San Diego, Calif. 92152, USA — Mrs. B. Webster
- Weih, Herta*, Dr., 1020 Wien II, Schüttelstrasse 19/6/15, Österreich
- Weinert, Herbert*, Dipl. Psychologe, 8212 Freital 5, Hirschbergstraße 6, DDR — Frau Lilli Weinert
- Wendler, Jürgen*, Dr., Facharzt HNO, 1055 Berlin, Nordmarkstr. 15, DDR
- Wenk, Reinhard*, Lehrer, Berlin, Schulzendorferstr. 22b, DDR
- Wepman, Joseph*, University of Chicago, Chicago, Illinois, USA
- Whitley, Eileen Marjorie*, Senior Lecturer in Phonetics, University of London W.C.1, Gower Street, Great Britain
- Wiede, Erwin*, Doz. Dr., Pädagogisches Institut, Güstrow, Bistede 47, DDR
- Wienold, Götz*, Dr., Wissenschaftlicher Assistent, 44 Münster (Westf.), Johannisstr. 12 bis 20, BRD
- Wierzchowska, Bożena*, Dr., Rakowiecka 22A/12, Warszawa 12, Poland
- Wierzchowski, Józef, M. A.*, Rakowiecka 22A/12, Warszawa 12, Poland
- Wiesinger, Peter*, Dr., Forschungsinstitut für deutsche Sprache — Deutscher Sprachatlas, 355 Marburg/Lahn, Kaffweg 3, BRD
- Winckel, Friedrich*, Prof. Dr. Ing., Berlin 12, Straße des 17. Juni 135
- Winston, Millicent*, Univ. Lecturer, Université Laval, Québec 10e, Canada
- Winter, Werner*, Prof. Dr., University of Kiel, BRD
- Witting, Claes*, Docent, Fonetiska institutionen, Thunbergsvägen 3 D, Uppsala, Sweden
- Wode, H.*, Dr., 78 Freiburg i. Br., Englisches Seminar der Universität, BRD — Frau Barbara Wode
- Wolter, Hugo*, Amanuensis, Fonetisk Institutt, Universitetet i Bergen, Bergen, Norwegen
- Worth, Dean*, Professor Dr., University of California, Los Angeles 24, California 90024, USA
- Wünschová, Felicitas*, Dr., Ústav pro jazyk český ČSAV, Praha 1, Letenská 4, Czechoslovakia
- Wurm, Stephen A.*, Professor, Dept. of Linguistics, Institute of Advanced Studies, A.N.U., Box 4, G.P.O., Canberra A.C.T., Australia
- Wurzel, Wolfgang Ullrich*, Strukturelle Grammatik, DAW, 108 Berlin, Otto-Nuschke-Str. 22/23, DDR
- Young, John*, Professor, Department of Asian and Pacific Languages, University of Hawaii, Honolulu, Hawaii 96822, USA
- Zabrocki, Ludwik*, Professor Dr., Poznań, Kanclerska 11, Poland
- Zacharias, Christina*, Dr., (Pädagogisches Institut), Erfurt, Richard-Breslau-Str. 6, DDR
- **Zacher, Oskar Ch.*, Professor, Irkutsk 11, ul. Transportnaja, d. 10, kv. 45, USSR
- Zagoruiko, Nikolaj G.*, kandidat techničeskich nauk, Institut matematiki SO AN SSSR, Novosibirsk 90, USSR
- Zatovkaňuk, M.*, Docent Dr., Praha 4-Spořilov, Hrusická 2511, Czechoslovakia
- Zichová, Marie*, Dr., Praha 10, Šrobárova 36, Czechoslovakia
- Zimnjaja, Irina Aleksejevna*, kandidat filologičeskich nauk, Moskva G 34, Metrostrojevskaja 38, USSR
- Zinder, L. R.*, Professor, Universita im. Ždanova, Leningrad, Universitetskaja naberežnaja, dom 7, podjezd 9, USSR
- Zlatoustova, Ljubov Vladimirovna*, Gerceev ul. 20, Laboratorija prikladnoj lingvistiki MGU, Moskva, USSR
- Zwirner, Eberhard*, Professor Dr. Dr., Deutsches Spracharchiv, 44 Münster/Westf., Stadtstraße 39, BRD — Frau Zwirner

Ахундов, А. А.	Фонематический состав аффиксальных морфем в тюркских языках	*
Akin, J.	The anglicization of Hispanic place names in Colorado	105
al-Ani, S. H.	Literary Arabic X-ray sound films	109
Altevogt, R.	Die vibratorische Kommunikation von <i>Uca tangeri</i> und <i>Uca inaequalis</i>	111
Annan, W. R. B.	A comparative study of the vowels of different accents of English	115
Антипова, А. М.	Интонационная структура восклицательных предложений в ее взаимодействии с их лексико-грамматической структурой	117
Antonelli, A.	The intelligibility of the sensitized speech tests in the disorders of the central auditory mechanisms	121
Arndorfer, H.	Das Phänomen der Lautbedeutsamkeit aus biophonetischer Sicht	125
Артемов, В. А.	Речевой акустический сигнал и его восприятие	129
Avram, A.	Sur le rôle de la fréquence dans la perception de l'accent en roumain	137
Багмут, А. И.	К характеристике интонации повествовательного предложения в славянских языках	141
Байчура, У.	Инструментально-фонетические данные относительно интонации и словесного ударения в тюркских языках	145
Bakis, R.		
—Rothauser, E. H.		
—Maiwald, D.	Über die Anwendbarkeit einfacher prosodischer Regeln für die Sprachsynthese	149
Bartoš, L.	Quelques observations sur le consonantisme de la modalité cubaine de l'espagnol	153
Барышникова, К. К.	Фразовое ударение как воспринимаемый и акустический компонент интонации	157
Baštecký, J.—Vinař, O.		
—Borovičková, B.		
—Maláč, V.		
—Zichová, M.	Speech errors on delayed auditory feedback	159
Becker, R.	Über das Erkennen synthetisch erzeugter vokalähnlicher Klänge durch normalsprechende, stammelnde und lese-rechtschreibschwache Schüler	163
von Beckerath, T.	Die Hysterie in der Sprechstimme	*
Bergsveinsson, S.	Die Pause im Kontext in freier Erzählung	167
Bierwisch, M.	Two critical problems in accent rules	171
Birnbaum, H.	Two problems of sound perception and phonemic interpretation in Polish	173
Black, J. W.	The magnitude of pitch inflections	177
Blanchard, I.	Phonetic patterns of consonant substitution	183
Блохина, Л. П., —Потапова, Р. К.	Идентификация речевого сигнала в зависимости от его звуковой структуры	185
Blom, J. G.	The application of information theory to vowel-recognition experiments	189
Bogdan, M.	Some notes on non-English sounds in continental English	197
Бондарко, Л. В.	Восприятие дифференциальных признаков и слоговая структура речи	201
Borkowska—Gaerstig, D.		
—Nowosielski, J.		
—Wierzchowska, B.	Significance of electroacoustic aids in development and perception of lingual signals in the young deaf child	205

Borovičková, B.	The relevant regions of Czech vowels from the perception point of view	209
Boudreault, M.	Les intervalles mélodiques	**
Brend, R.	Acoustical support for African field work	211
Брока, В. А.	О некоторых принципах сопоставления речевой интонации	*
Бровченко, Т. А.	Энергетические корреляты словесного ударения	215
Брызгунова, Е. А.	Основные типы интонационных конструкций в русском языке и их восприятие	219
Bystrzanowska, T.		
—Dagajev, B.	L'influence de la fatigue sur la discrimination de la voix filtrée	225
Calero, C.	Aperçu général sur la perception du langage à redondance réduite	231
Carney, E.	Patterns of auditory resemblance in English consonants	**
Carnochan, J.	The vowels in Bachama	235
Carton, F.	Pente et rupture mélodique en français	237
van Caspel, P.	Acoustic aspects of poetry	243
Серлитис, Л.	Использование свободной речи в фонетических исследованиях	247
Cerdà, R.	Méthodologie de la mesure physiologique d'après les films radiologiques	*
Хачатрян, А. А.	К вопросу об артикуляционных особенностях аффрикат армянского языка	249
Charbonneau, R.	Le phonème (ε) en français canadien	253
Christov, Ph.	A statistical approach to speech analysis by waveform	265
Cohuț, C.		
—Mărdărescu, M.	Les voyelles mégléno-roumaines	271
Constant, G.	L'appareillage auditif pour enfants sourds	*
Daan, Jo	Diachronische Betrachtungen über das Lautsystem der niederländischen Sprache nach 1600	277
Dillon, M.	An example of phonetic mixture and semantic fusion	**
Domanski, R.		
—Dzioch, E.	Die Behandlung des Stotterns mittels Echo-Apparat	283
Дубовский, Ю. А.	Интонация встречного вопроса в современном английском языке	*
Dukiewicz, L.	Die Aussprache der ę, ą, — Laute in der gegenwärtigen polnischen Sprache	287
Дворжецкая, М. П.	Интонация перечисления в основных коммуникативных типах предложения	*
Джаперидзе, З. Н.	Некоторые закономерности восприятия звуковой формы речевого потока	291
Elekfi, L.	Beiträge zur Prädiktabilität des satzbeginnenden Akzents im Ungarischen	293
Eras, H.	Die Entwicklung eines Tonhöschreibers für phonetische Forschungen	297
Eustace, S.	Present changes in English pronunciation	303
Ezawa, K.	Sprachnorm und Sprechnorm	307
Файн, В. С.	О непрерывно-групповом подходе к опознаванию речевых образцов	*
Fasske, H.	Zur Rezeption und phonologischen Interpretation des *o nach Labialen und Velaren im Niedersorbischen	313
Faure, G.	La description phonologique des systèmes prosodiques	317
Filip, M.	Some aspects of high-accuracy analog fundamental-frequency recording	319
Fintoft, K.	„Toneme 3“ in Norwegian	323
Firbas, J.	Note on the prosodic features as means of functional sentence perspective	327
Fliflet, A. L.	Norwegisch-schwedische phonetische Kennzeichen morphologischen Unterschiedes zwischen „identischen“ Lautverbindungen	**

Flint, E. H.	Intra-syllabic pitch movement in communicative Australian English utterance morphemes	331
Frinta, A.	Quelques remarques méthodologiques	335
Fry, D. B.	Reaction-time experiments in the study of speech processing	337
Fujisaki, H.		
—Kawashima, T.	Roles of pitch and higher formants in perception of vowels	347
Гайдучик, С. М.	Влияние фразового ударения на длительность немецких гласных *	*
Гаприндашвили, Ш. Г.	Вопросы теории дифференциальных признаков	351
Garvin, P. L.		
—Worthy, R. M.	The diagnostic function of idiosyncretic phonetics	357
Geißner, H.	Rhetorische Analytik	361
Gendron, J.—D.	Etude expérimentale des aspirées sonores en malayalam	365
Генина, В. Г.	Интонационная структура предложения с абсолютной конструкцией в английском языке	*
George, D.	Phoneme cognition and perception	371
Golopenția, —Eretescu, S.	La perception acoustique du vers	**
Goosens, J.	Ist die These „Lautwandel durch innere Kausalität“ sprachgeographisch beweisbar?	375
Göschel, J.	Formantbewegungen in Vokalen deutscher Mundarten des westlichen Harzes	381
Гожин, Г. М.	Консонантные дифтонги [či], [gʲi] в молдавских говорах	391
De Grève, M.	Phonétique et structuralisme	395
Grobben, L. M.	Transient phenomena in music and speech	399
Gsell, R.	Intonation en français**	405
Hadj-Salah, A.	L'impulsion syllabique et sa perception d'après les anciens phonéticiens arabes	407
Hammarström, G.	Pseudo-units in phonetics	411
Hamp, E. P.	Systems of lateral sounds and perception	415
Handzel, L.	Musische Faktoren in der Sprache der tauben Kinder	419
Handzel, L.		
—Jarosz, A.	Über die Hörbarkeit der Sprachmelodie	421
Harms, L. S.	Interchange in speech communication	425
't Hart, J.		
—Cohen, A.	Experiments with artificial intonation contours	429
Heike, G.	Lautdauer als Merkmal der wahrgenommenen Quantität und Betonung im Deutschen	433
Herbst, L.	Über Faktoren, welche die Perception der individuellen Hauptsprechtonhöhe fördern oder hemmen	439
Hořejší, V.	L'utilisation de la chronologie relative dans la description structurale de l'évolution phonologique des langues	445
Huckleberry, W. A.	The disappearing postvocalic [r] of „General American“ English speech	449
× Piescu, M.	La fréquence des finales vocaliques dans les parlers rhétoromans	453
Janota, P.	Perception of stress by Czech listeners	457
Kaiser, L.	Information by tones and single vowels	463
Кандинский, Б. С.	Интонация и контекст	*
Каспарова, М. Г.	К вопросу о членении речи при восприятии	469
van Katwijk, A.		
—'t Hart, J.	Intelligibility of syllable-tied interrupted speech	471

Kaulhausen, M. H.	Über den Ursprung der Sprachmelodie in der klassischen und modernen deutschen Lyrik	475
Kelly, J.	Phonological formulae for the Polish verb	**
Keramitchievska, L.		
—Keramitchievski, S.	Linguistic determination of the intelligibility of the speech sounds	479
Keramitchievski, S.	Some basic components of perception of speech sounds	485
Kinsbourne, M.	Speech perception in aphasia	**
Klajman, S.	Some physiological speech problems in hiperbaria	489
Koch, W. A.	Phonologie und Textanalyse	493
Kohler, K. J.	Vowel, consonant and syllable re-examined	499
Козьмин, О. Г.	Интонация побудительных предложений в немецком языке	*
Krech, E.—M.	Perzeptionsuntersuchungen zur Relevanz des coup de glotte im Deutschen	503
Kress, B.	Zur Einpassung anglo-amerikanischer Wörter in das Isländische	507
Kurka, E.—Stock, E.	Zur Beeinflussung der Prosodie in der Sprache von Stotterern durch simultane Musikperzeption	511
Lane, H.	On the necessity of distinguishing between speaking and listening	515
Langová, J.		
—Morávek, M.		
—Novák, A.		
—Petřík, M.	Der Leeeffekt vom Gesichtspunkt der Informationstheorie	519
Laver, J.	The synthesis of components in voice quality	523
Lawrence, T. Z.	Certain phonetic tendencies perceived in the idiolects of selected native Texans	527
Lebrun, Y.	Clinical evidence against the motor theory of speech perception	531
Lee W. R.	Noticing word-boundaries: a brief investigation	535
× Lehiste, I.	Diphthongs versus vowel sequences in Estonian	539
Lejska, V.	Einfluss der Mundart auf die Artikulation der Kinder	545
Liiv, G.	Some notes on the experimental study of spectral parameters of vowel-consonant transitions in Estonian	549
Lindner, G.	Die Beeinflussung der Beurteilung isolierter synthetischer Vokale durch Sukzessivkontrast	553
Liška, J.	Length of vowels in Sotak dialects	559
Lisker, L.		
—Abramson, A. S.	The voicing dimension: some experiments in comparative phonetics	563
Abramson, A. S.		
—Lisker, L.	Discriminability along the voicing continuum: cross-language test	569
Lotzmann, G.	Künstlerische Komposition und stimmlicher Ausdruck bei Bertold Brecht	575
Lüdtke, H.	Die Alphabetschrift und das Problem der Lautsegmentierung	579
Луцихина, И. М.	Аудирование речевых сообщений с точки зрения их структуры и релевантности для оператора	585
Махмудов, А.	ЛабIALIZED фонемы в современном узбекском языке	589
Mahnken, I.		
—Matešić, J.	Akzentoppositionen in den serbokroatischen Dialekten: Zum 5-Akzent-system der slawonischen Mundarten des Štokavischen	593
Maláč, V.	Some perceptual cues of Czech consonants	597
Malmberg, B.	De Ferdinand de Saussure à Roman Jakobson — l'arbitraire du signe et la substance phonique du langage	599
× Mańczak, W.	Deux lois de l'évolution phonétique	605

Maneca, C.	Considérations statistiques sur les finales vocaliques en roumain . . .	609
Marshall, J. C.	The influence of syntactic structure on the perception of intonation . . .	**
Martens, P.	Optische Unterstützung akustischer Sprachsignale	613
Massa, A.		
—Stevens, D. M.		
—Takefuta, Y.	Quantification of signals in the intonations of American English by human listeners	617
Matějka, L.	Czech quantity in paradigmatic and syntagmatic procedures	621
McConnell, F.		
—Robertson, J. B.	Auditory perceptual skills of culturally disadvantaged children	625
McDavid, R. I. jr.	Some problems of over-all patterning	631
Мегрелишвили, Р.	Спектральный анализ шелевых согласных французского языка в сочетании с гласными в позиции СГ	*
Mehnert, D.	Einrichtungen zur Mehrfachanalyse akustischer Signale	633
Meier, G. F.		
—Kyritz, J.	Einige Kriterien für die Erkennung akustischer Signale	637
Meinhold, G.	Nasale und orale Vokale — Struktur und Perzeption	641
Meinsma, G. L.	A scaling technique in an experiment with vowel-like sounds	645
Мельников, Г.	Усовершенствование „треугольника“ международного фонетического алфавита для отражения степени акустической противопоставленности гласных	*
Merlingen, W.	Was nimmt man bei den Sprachlauten wahr?	651
Микалаускайте, Е.		
—Сведевичус, Б.	Релевантные признаки просодии слога в современном литовском литературном языке	655
—Пакерис, А.		
Mikus, R. F.	Analyse quantitative des cas difficiles	659
Mol, H.	A model of the vowel system of Dutch	661
Нейланд, М. А.	Сопоставительный анализ акустических свойств английских и латышских дифтонгов	667
Neustupný, J.	Perception and linguistic vagueness	**
Nickel, G.	Some suprasegmental criteria for contextual relations between sentences in English	671
Novák, A.—Morávek, M.		
—Petřík, M.		
—Langová, J.	Einfluss des Lee-Effektes und des Geräusches auf die Sprache	675
Novák, L.	Définition de la syllabe et de ses composantes (voyelle, diphtongue, sonore, consonnes)	679
Novotná, J.	The acoustic features of the positional variants of nasals in Czech	681
Нурмаханова, А. И.	Некоторые вопросы исследования интонации типов предложений в тюркских языках	685
Ohman, S.—Persson, A.		
—Leanderson, R.	Speech production at the neuro-muscular level	689
Ohnesorg, K.	En marge de quelques problèmes pédophonétiques	697
Oliverius, Z. F.	Phonemic syncretism and the process of communication	701
Ondráček, J.	Some notes on vowels and geminated consonants in contemporary Italian	705
Ondráčková, J.		
—Lindqvist, J.		
—Fant, G.	Electrical glottography	709

Onishi, M.	Perception of quality against quantity	713
Pahn, J.—		
—Seifert, Ch.	Anaesthesiebehandlung bei Dys- und Paraesthesien im Kehlkopfbereich	717
Palková, Z.	Einige Probleme des Prosarhythmus	719
Панкрац, Г. Я.	Средства и способы выделения коммуникативного ядра предложения	*
Paulsen, K.	Zur phylogenetischen Entwicklung des Kehlkopfskelettes von den Amphibien bis zum Menschen	723
Paulsen, K.	Stimmbildung in der aufsteigenden Wirbeltierreihe	727
Penzl, H.	Zur Lautwiedergabe in Lehnwörtern des Althochdeutschen	731
Petitclerc, G.	Reducing and perseveration of immature timing between an acoustic speech signal and its perception	735
Pignataro, O.		
—Teatini, G. P.	Sur la probabilité de détection („detectability“) de différents messages vocaux	737
Плакунова, Т. Е.	Сonorные согласные литовского языка в смешанном дифтонге	*
Pražić, M.	Artificial distortion of words	741
Pulgram, E.	Phonetic signals of syllable boundaries	747
von Raffler		
Engel, W.	The auditory perception of distinctive features in a six year old child	751
Ревтова, Л. Д.	Интонация обращения	755
Richter, H.	Intersubjektive Übereinstimmung von Intonationsbeurteilungen	759
Rigault, A.	L'assimilation consonantique de sonorité en français: étude acoustique et perceptuelle	763
Rischel, J.	Acoustic features of syllabicity in Danish	767
Roceric		
—Alexandrescu, A.	Intonation and predictability	771
Rodón, E.	A phonetic realization of the past morpheme in English	773
Romportl, M.	Vibranten als Phonemkategorie	775
Rossi, M.	Sur la hiérarchie des paramètres de l'accent	779
Румянцев, М. К.	Время звучания тонированного слога (на материале китайского языка)	*
Šabršula, J.	La description phonétique de la négation française et la stylistique	787
Saltarelli, M.	Intonation as a function of stress and pitch grammatical cycles: Italian	**
Шаумян, К. С.	Фонологическая система как единство комплементарных диакритических систем	791
Schädlich, H. J.		
—Eras, H.	Vergleichende Untersuchungen über die Satzintonation in deutschen Dialekten	793
Schindler, O.	L'enregistrement tridimensionnel de l'objet sonore	797
Schneider, A.	Suvag-lingua und die phonetische Korrektion	799
Schwanzer, V.	Charakteristische schallnachahmende Lautgruppen im Deutschen	803
Scully, C.	Vowel quality and wave shape	807
Sedláček, K.		
—Sychra, A.	Relations between the acoustic, articulatory and psychological parameters of the emotional expressions in speech	811
Sedláčková, E.	Differenzierung der Formantengebiete in den Anfängen der Stimm-entwicklung des Kindes	813
Shen Yao	An acoustical signal in Mandarin (Chinese) questions with interrogatives	817

Silva, G.			
—Pratt, B.	Phonemes: An algorithm for the transcription of French text into phonetic symbols by a computer	821	
Singh, S.	Interrelationship of English consonants	825	
Skalička, V.	Die Typologie der Lautsysteme	829	
Skaličková, A.	The English Diphthongs	833	
Slama—Cazacu, T.	Quelques principes psychologiques et psycholinguistiques dans l'étude de la perception de la parole	837	
Slis, I. H.	What causes the voiced-voiceless distinction?	841	
Söderberg, G.—G.	Le phonétique et l'audiologie	845	
Sovák, M.	The acoustic signal and the reflex theory	849	
Sovijärvi, A.	On the transitions in the light of X-ray films	851	
Sovijärvi, A.	Examples of X-ray films from Finnish	857	
Stelzig, H.	Die Bedeutung der auditiven Kontrollfunktion für die Steuerung von Sprechsituationen als typischen Informationseinheiten	859	
Stewart, J. M.	A theory of the origin of Akan vowel harmony	863	
Stock, E.	Zur Bewertung des Tonhöhenverlaufs im Aussagesatz	867	
Šupáček, I.	Speech intelligibility of the deaf and hard of hearing persons	871	
Șuteu, V.	Sur le „locus“ des consonnes roumaines	875	
Takefuta, Y.	A study of relative efficiency of acoustic parameters in the intonational signals of American English	879	
Тарановский, К. Ф.	Звуковая фактура стиха и ее восприятие	883	
Tätaru, A.	Unusual mistakes in hearing and pronouncing foreign sound-sequences	887	
Tatham, M. A. A.	Specifying speech	889	
Teatini, G. P.	Perception of low-redundancy speech in children	893	
Tembrock, G.	Motivation und Information bei der Lautgebung von Säugetieren	897	
Tervoort, B. Th.	Some aspects of speech of deaf children	901	
Theiner, Ch.	Zur phonematischen Differenzierungsfähigkeit der Vorschulkinder mit funktionell bedingtem Stammeln	903	
Tillmann, H. G.	Kognitive Voraussetzungen der Theorie einer generativen Signalphonetik	907	
Томанович, В.	Об экспрессии звуков	911	
Toporišič, J.	Relevanz der Gestaltelemente der slovenischen Toneme	913	
Торсуева, И. Г.	Акустические характеристики смыслового членения предложения	917	
Trim, J.	Cues to the recognition of some linguistic features of whispered speech in English	919	
Trojan, F.	Bericht über die Ergebnisse der Enquête zu Jakobson's Gesetz vom Schichtenbau des Sprachlautsystems	925	
Trojan, F.	Die statistischen Unterlagen der Sprachlautcharakterologie	931	
Tronczyńska, J.	Manometric evaluation of nasality in cleft palate speech	935	
Trost, P.	Vokalquantität und Silbe im Deutschen	939	
Truby, H. M.	The perception and common misperception of infant pre-speech	943	
Tscheschner, W.	Über die Gestaltkriterien akustischer Sprachsignale	949	
Туркбенбаев, И.	Интонация простых вопросительных предложений в казахском языке	*	
Ulbrich, H.	Zur auditiven Interpretation von deutschen (r)-Allophonen	955	
Урбелене, Я.	К вопросу о редукции гласных в современном литовском языке	959	
Vachek, J.	Some remarks on „juncture“ in phonological analysis	963	
Valdman, A.	Nasalization in Creole French	967	
Vallancien, B.—			
—Djourno, A.			
—Martin, A.			
—Faulhaber, B.	Les perspectives de la logographie	973	
Vasilii, E.	Rules ordering and evolution of phonemic systems	979	
Васильев, В. А.	Акустические компоненты речевого сигнала и их функции	981	
Венцолл, М.	Фонемная структура слова в современном английском и латышском языках	985	
Вербицкая, Л. А.	Акустическая характеристика воспринимаемых звуковых единиц русской речи	989	*
Вишневская, Л. Я.	Интонация повествования и вопроса в языке хаусса	*	
Власова, З. С.	Коммуникативно-интонационное членение предложения в немецком языке	*	
Vrtička, K.	The relation between the objective frequency of the fundamental tone and the subjectively perceived pitch and melody of the oesophageal voice	993	
Wajskop, M.	Identification de voyelles en fonction de leur durée	997	
Webster, J. C.			
—Carpenter, A.			
—Woodhead M.	Perceiving complex speech and non-speech sounds	1001	
Weihls, H.	Der stimmliche Ausdruck in Sprechstimme und Musik	1005	
Wendler, J.	Zur auditiven Steuerung der Sprechintonation	1009	
Whitley, E. M.	Some problems of quantity in certain Indian languages**	1015	
Wiede, E.	Einige Bemerkungen zur Artikulationsbasis des Russischen und Deutschen	1017	
Wienold, G.	On phonological ambiguity	1019	
Wierzchowska, B.	Aus den kinoröntgenographischen Untersuchungen der polnischen Verschlusslaute	1023	
Wierzchowski, J.	Laut, Bedeutung und Sprachgrundeinheit	1025	
Wiesinger, P.	Die Bedeutung mundartlicher Aussprachen durch Mundartsprecher im Deutschen	1027	
Winckel, F.	Darstellung des Sprechverhaltens als statistische Tonhöhen und Formantverteilung mittels Langzeitanalyse	1031	
Wode, H.	Sprechstörungen bei Parkinsonpatienten	1037	
Wurm, S. A.	Some phonetic phenomena in New Guinea languages	1041	
Wurzel, W. U.	Regeln für die Fremdwörterbetonung im Deutschen	1043	
Zabrocki, L.	Kybernetische Sprechmodelle	1047	
Zacharias, Ch.	Form—Funktionsbestimmung der Intonation einer Interferenzvariante des Fragesatzes	1051	
Zacher, O.	Zur phonologischen Wertung der Intonationsmittel	1055	
Загоруйко, Н. Г.	Проблема распознавания речевых сигналов как сложная система	1057	
Zichová, M.	Quelques remarques concernant la durée relative des voyelles du français contemporain	1063	
Зимняя, И. А.	О восприятии речевой интонации	1067	
Зиндер, Л. Р.	Фонема и восприятие	1071	
Златоустова, Л. В.	Временные характеристики русской речи	1075	
LA RÉUNION DU CONSEIL PERMANENT			1079
REPORT OF THE GENERAL MEETING OF THE ISPhS			1081

BY-LAWS PROVISIONALLY ACCEPTED AT THE GEN. MEETING OF THE ISPLS	1083
DIE SITZUNG DER INTERNATIONALEN VEREINIGUNG FÜR BIOPHONETIK	1087
RÉUNION DE L'ASSOCIATION INTERNATIONALE POUR LES ÉTUDES PHONO-	
LOGIQUES	1089
COMPTE RENDU DE LA SÉANCE DU COLLÉGIUM DE PHONOLOGIE EXPÉRI-	
MENTALE	1091
COMPTE RENDU DE LA RÉUNION DE I. A. L. S. A.	1093
CLOSING SESSION (<i>Prof. D. B. Fry</i>)	1095
LIST OF PARTICIPANTS	1097

* Published in *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, Akademie-Verlag, Berlin.

** Text of paper not submitted.