LINGUISTIC FACTORS IN SPEECH PERCEPTION

A.S. STERN

Dept. of Phonetics
Leningrad State Univ.
Leningrad, USSR, 199034

ABSTRACT

The hypothesis tested in this research is that certain linguistic characteristics have a material influence on speech perception. A statistical model based on analysis of variance in perceptual data is proposed, where significant factors are assumed to be the perception cues and their levels to be decision making units. The investigation of the model has enabled us to elucidate a number of psycholinguistic features of the speech perception process, the typological properties of a given language as well as some characteristics of perceptive ability development in both native language acquisition and second-language learning.

HYPOTHESIS, METHODS, MATERIAL

In the present work the perception of cardinal psycholinguistic units, i.e., syllables, words, sentences and texts, was studied. Listening to speech stimuli was chosen as an experimental procedure, since it seems to be a perceptual activity that is mainly dependent on the processing of sound sequences and is not closely related to the higher levels of speech comprehension. A group of 7-10 subjects was asked to listen to sets of speech stimuli presented against the background of some distortion and to write them down. The texts were presented several times, while other stimuli only once. Different kinds of distortions or their combinations were used: a) objective distortions (white noise, distant reception, synthetic speech stimuli, accented speech), and b) subjective (poor hearing, poor knowledge of the language, aphasia). The quantitative aspect of distortion namely, the signal/noise ratio (S/N), the degree of hearing loss, the level of performance in the second language, etc. was also varied.
Each speech segment can be described on the basis of its correct perception frequency. Besides, one may obtain a number of ratings for various linguistic features. For example, the word "ruka" (hand) is a noum (a level of the factor Parts of Speech), with the highest possible frequency of occurrence (a level of factor F_{ob}), containing the stressed "а",bisyllabic, etc. Correct recognition of the word "ruka" is assumed to be determined by these factors, or more precisely, by their levels. Hence, it is quite natural to use analysis of variance to discover the significant linguistic features (factors) and to establish a hierarchy among them. Results of this analysis have yielded a statistical descriptive model of speech segment perception.
Let us consider a fragment of such a model, giving the correlation ratio η_x^2 of some factors in word recognition: 1 - against the background of white noise at S/N = -6dB; 2 a,b - in hard of hearing adults with different degrees of hearing loss; 3 - for German students who perceived Russian words in white noise at S/N = = -2dB. The significant factors are underlined (see the Table).

Table

Factors \ Experiments	1	2a	2b	3
Stressed Vowel	0.052	0.020	0.020	0.006
Voiced/Voiceless	0.000	0.007	0.002	0.005
Soft/hard	0.017	0.015	0.004	0.009
Length in Syllables	0.073	0.010	0.006	0.013
Parts of Speech	0.018	0.040	0.030	0.094
F_{ob}	0.012	0.003	0.002	0.043

For correct use of analysis of variance, the factors being investigated in the experimental material should be orthogonal. In most cases balanced articulatory tables were used /2/.
Our conclusions are based on the analysis of about 50 experiments, giving approximately 70,000 responses. These experiments were conducted, in part, in collaboration with my collegues. The study of the models obtained has made it possible to discuss three groups of problems.

I. THE PSYCHOLINGUISTIC FACTORS IN SPEECH PERCEPTION

A. Isomorphism of Models for Speech Unit Perception at Different Linguistic Levels in Auditory Listening Tests.

This is confirmed, first, by the fact that the models for speech unit perception at all linguistic levels are shown to be analogous, and, secondly, the same factors hold for units of different levels. For example, the factors Stressed Vowel and some Distinctive features of consonants are significant for both syllable and word recognition. Thus, a certain isomorphism of linguistic levels in the process of listening may be postulated. It should be noted that the obtained factors act simultaneously in every instance and no "input" can be found into a set of this type.

B. Similarity in Mechanisms of Perception Irrespective of the Distortion Type.

Each type of distortion is characterized by an individual set of factors or a hierarchy of these factors. There are, however, factors which turn out to be significant in the majority of cases. Among them we find the following: relative frequency of occurrence and length in syllables for words, the stressed vowel, parts of speech. To conclude, it should be mentioned that there is an evident similarity in the mechanisms of speech perception under different conditions of distortion, which not only justifies the accepted approach towards speech pathology, insufficient knowledge of the language and noise as a distortion, no matter what its nature may be, but also helps to understand every single case on the basis of distortions of other types.

C. Differences in Mechanism Depending on the Degree of Distortion.

When the type of distortion is constant but the degree is altered not only common but specific factors as well are revealed, besides, their ranks may vary. For example, F_{ob} of speech units (syllables or words) is found to be one of the most important factors in poor reception conditions and to decrease in significance as the reception conditions improve. The factor Parts of Speech is insignificant under poor reception conditions whereas under superior conditions it becomes a factor of great value. Thus, it can be said that the analysis revealed both common and specific features. The first of these two findings, i.e. the existence of common features, was not unexpected. The second one, on the other hand, is difficult to predict and, therefore, is mostly ignored by researchers. In order to sum up the results of this section and of the preceding one, let us underline that the common features in mechanisms of perception are at work in all types of distortion, whereas specific features depend on the degree of distortion.

D. An Extension of Jakobson's Regression Hypothesis.

Let us now look at the data from a different angle. R. Jakobson proposed a hypothesis according to which aphasic speech disorders mirror the process of language acquisition in children. The data on the factor levels indicate the following: vowels are better recognized than consonants, /a/ is much more easily recognized than /ɑ/; choreic words are easier than iambic ones; nominative case is better perceived than other cases; the direct object is superior to the indirect object in the number of correct responses. The active construction is recognized more easily than the passive one. The dialogue is easier to perceive than the monologue, words of frequent occurrence are recognized correctly more often than rare words. Is is clear that the first members of the oppositions are acquired earlier in the ontogenesis. We can therefore attempt to extend Jacobson's hypothesis in the following way: the linguistic features which are the earliest to have been acquired are the most stable in all types of distortion.

E. The Existence of Simple and Complex Factors Functioning as One Whole.

Some of the factors are simple and cannot be further disintegrated into other features (i.e. distinctive features of the phonemes or parts of speech). Other factors, such as syllabic contrast or communicative type of text, may be conceived as a combination of more elementary features. But in the process of speech perception these complex features may become crucial, that is, they function as a whole. An increase in the weight of such complex features is often caused by an improvement in reception conditions. This fact is in agreement with some recent psychological investigations.

F. Differences in the Perception of Isolated Units and Units in Context.

Comparison of sets of significant factors for isolated words and words included in a text indicates that some of them are present in both test conditions. For most factors, however, a decrease in significance or a complete loss of significance is observed. Thus, the mechanism of perception is different for isolated words and words in context.

G. Simultaneous Perception of Speech Unit as a Whole and in Elements.

Some factors are related to elements into which the speech units can be subdivided (e.g. stressed vowels), whereas the others describe the unit as a whole (e.g. the rhythmic structure, frequency of occurrence). Since both types are significant simultaneously, one may suppose that the recognition of the whole unit and that of its parts occurs parallelly. Let us consider some additional facts. If we compare the hierarchies of all factors for words and syllables under similar conditions we can clearly see that for S/N = = -6dB rank test ρ is +0.86, for 0dB it is +0.60, and at +4dB it is +0.09. These data indicate that under poor reception conditions the mechanism of phonetic processing of a word is highly efficient which is not the case under good reception conditions. In another experiment listeners were given words spoken by non-native speakers of Russian (the Agul) and parts of these words pronounced with a strong accent. It was found that ρ (rank test) for the correct recognition of words and their parts in 4 different groups of listeners varied from -0.10 to +0.17, that is, there was actually no correlation at all. This signifies that words were perceived regardless of the presence of some distorted segments, i.e. as whole units.
Moreover, when German students recognized Russian words both masked and not masked by noise, correct recognition scores in the latter case were twice as high as in the former case. This improvement was due to perception of both familiar and infamiliar words. Thus, a possibility of phonemic decoding has been demonstrated. Now we can amend the rule as follows: speech units are perceived simultaneously as sequences of elements and as integral units (Gestalt), the strategy depending on the perceptual situation.

H. Simultaneous Involvement of All Linguistic Levels Regardless of the Type of the Unit to be Perceived.

To make this item clear, let us take our data on words. Word perception is determined by the following factors: certain distinctive features of consonants and vowels (the sound level), length of words in syllables (the syllabic level), part of speech and length in morphemes (morphemic level), the number of quasiomonymes (word level) and F_{ob} (the text level). This indicates that various linguistic levels are involved in the perception of speech units at the same time.

I. Speech Perception as an Action.

It is generally considered that the probability prediction is based on the fact that the listener is an active recipient of speech. Our experiments have confirmed the significance of the probability factor. Thus, the greater the probability of a word or syllable, the higher the correct recognition scores. An additional experiment has shown, however, that this mechanism is closely related to the frequency distribution in a sample,i.e. when frequencies of elements correspond to their linguistic probabilities this dependence is the lowest. Conversely, when the elements are equally distributed the direct dependence is higher.When the distribution is reverse, i.e., when elements with high probabilities occur rarely and vice versa, the dependence is also higher, but the correlation will have an opposite sign ("-") indicating that high probability elements are harder to recognize than low probability ones. Thus, the active character of perceptual processes is revealed in an interplay of the listener's sociolinguistic experience and the current analysis of frequency distributions in a given sample. The listener's activity is also revealed in series of choices he has to make: of a perceptual (phonetic) base from those he has at his disposal; of a morpheme from a corresponding morphemic class; of a word from a set of similar words, etc. All this applies only to speech units (from sounds to words) presented in isolation. In a text,however, the role of this factor considerably decreases. On the other hand, a key word prediction factor emerges, whose activity is linked with the work of association mechanism.

II. THE PSYCHOLINGUISTIC TYPOLOGY OF LANGUAGES.

Comparison of significant factors for a number of languages, namely,Russian, German, English and French enabled us to obtain both universal and language specific factors. F_{ob} and Parts of Speech are examples of universal factors.Specific factors for the Russian language are the location of the word stress and word order. The former is non-existent in French while the latter in German. The word-length factor may serve as another example. In Russian, the word length in syllables is quite significant whereas word length in morphemes is of less value (η^2 is 2 times less). In German the situation is the reverse,word length in syllables being completely insignificant and word length in morphemes is in the forefront of significant factors. This latter fact is evidently connected with the greater "syntaxicality" of the German word. A projected analysis of other languages will help to establish a typology of languages at the perceptual level.

III. THE FORMATION OF THE PERCEPTUAL MECHANISM IN SPEECH ACQUISITION AND IN SECOND-LANGUAGE LEARNING.

A. A comparison of speech perception mechanisms in normal adults against the background of white noise,in hard-of-hearing adults, in normal children listening to speech in white noise and in hard-of-hearing children has shown that there was a +0.11 and a +0.14 rank correlation between adults and children for the same distortion type, and ρ = +0.50 between the two groups of children as well as the two groups of adults. This indicates that speech perception is determined by the age of the listener. It is especially important for children.

B. The Sets of factors and their hierarchy change in the course of second language learning, the degree of similarity with the native language mechanism decreases as that of the second language increases. For example, in the group of German students that participated in recognition tests of Russian words in white noise in their 1st, 3rd and 5th years at the university, ρ varied as follows: 0.40 → 0.28 → 0.18 as compared to the mechanism in German and 0.45→0.42→0.71 as compared to that in Russian.
On the basis of the above presented data it may be concluded that significant linguistic factors are perceptual cues (in the sense of the word introduced by S.Vygotsky and A.A.Leontyev), reflecting the elementary psychological operations of the speech perception processes.Moreover, the investigation suggests that the significance (and the maximum η^2_x) cannot be obtained unless an adequate way is found of determining factor levels (see the example on word length in Russian and German given above). The listener is assumed to make use of linguistic factors "keeping in mind" a particular level of factors. Hence, levels of linguistic factors are decision-making units.

Acknowledgements

The author wishes to thank Prof.L.R.Zinder for helpful comments on earlier versions of this report and his interset in all stages of the research. Thanks are due to my colleagues who took part in various experiments of the study. I also thank I.Panasjuk for looking through the English version of the text and making useful suggestions.

REFERENCES

1. L. Zinder, A. Stern. Factors Affecting Word Recognition. - Recent Trends in Soviet Psycholinguistics. New York, 1977-1978, p. 123-130.
2. A. Stern. Articulatory Tables for the Development of Perceptual Skills and Testing the Auditory Function. Leningrad State University Press, 1984 (in Russian).