SOME SOCIOLINGUISTIC CHARACTERISTICS OF PHONETIC ANALYSIS*

JOHN N.H. PELLOWE, G. NIXON AND V. McNEANY

There are numerous applications where it is important to know exactly what we mean when we say that the written value for an auditorily perceived segment of a speech signal IS such and such. One of these applications is research into language variation and change. There is little material on the reliability of transcriptions (except, especially, Ladefoged 1960). For the Tyneside Linguistic Survey (T.L.S.), of which this research is a part, the matter is of some importance, since we are concerned to develop and operate a general methodology for determining the ecology of varieties of spoken (primarily urban) English. That is, we identify speech varieties, determine their relative commonness or rarity, and define their distribution across social attributes. Before examining particular characteristics of phonetic variability, we shall need to sketch our methods of dynamically modelling linguistic variation. We need a dynamic or heuristically programmed model — which produces constructs, or results, rather than THE result — because of the complete ignorance of the underlying mathematical properties of language variation. There is an increasing amount of evidence, which space here forbids us to refer to (see Pellowe *et al.* 1971), which shows that when hearers perceive speech, they derive much information from the linguistic signal about their interlocutors, in addition to decoding in the strict sense. There are good reasons for assuming that this extra-linguistic information is derived from a function of dissimilarity between the acoustic signal and the hearer's linguistic experience as a speaker.

The hearer's assessment of hearer-speaker dissimilarity implies distance as a spatial metaphor, and our model exploits this metaphor. If we conceive of any particular hearer as a fixed point in space, then we may base that fixed point upon whatever it is which is used as the basis for comparison, and assume that his perception of the distance of other speakers from himself is a function of their dissimilarity. This spatial view seems capable of accounting for important contingencies: it underlines the nonbiuniqueness of the relations between social and linguistic variables; it allows for the different perceptions of speakers by hearers on the basis of perspective principles.

* Grateful acknowledgement of financial assistance is made to S.S.C.R. (U.K.)

The basis for our dynamic modelling of sociolinguistic variation is a principled search for a general optimum 'space', variety space (VSp), which will represent our data. The VSp will be a classification of the most general and natural kind. The VSp is a multidimensional space, each of whose dimensions with its scale is a CRITERION with its variants. A criterion is any feature of speech showing at least two variants across the population under consideration, and not LOGICALLY predetermined by the nature of any other criterion in the set of criteria currently in use. Any particular speaker will thus have a unique multicoordinate position in this space, and our ecological aim can be re-expressed as needing to know how speakers 'fill' this space. Our remarks about hearers imply that we expect the dispersion of speakers in the VSp to be 'clumpy', or discontinuous to varying degrees.

If we assume that each multicoordinate point in VSp is a variety, then the clumps or clusters in VSp, which can be viewed as swarms of highly similar Vs, are variety clusters (VC). In terms of what we have said about hearers, we next disperse the same sample of speakers who are in the VSp, into another,independent, multi-dimensional space on social attributes; we then seek a function which expresses the group properties of a VC in successfully predicting the social cluster (SC) allocation of those VC members. There are many other properties of the space which we are investigating but we cannot go into them here.

The crucial definition for our purpose now is that of an analytic variety: an $_{a}$V is A PROFILE OF LINGUISTIC VARIANTS WHICH EXHAUSTIVELY PLACES ANY SAMPLE OF SPEECH OF A PARTICULAR SPEAKER IN THE VARIETY SPACE AS DEFINED AT THAT MOMENT BY ITS CONTENTS AND DIMENSIONS. The qualification 'exhaustively' is important since it emphasises that the usefulness of any current version of the VSp is its dimensionality or criterial properties. Having said this, we need to examine the methods used to arrive at the particular dimensionality and particular contents of the space — namely the selection of criteria and the coding of informants as they result from analysis. What is predicted by the model itself (Pellowe *et al.* 1971) is that different analysts will select different criteria as a result of differences in the analysts' own positions in the space. Furthermore, an analyst's willingness to incorporate some criterion must depend on his exposure to it. We may characterise this limitation on exhaustiveness in terms of the geometrical properties of the analytical VSp. Insofar as a given selector under- or over- represents any delimitable subsets of possible criteria, he will be operating with a topologically deformed version of the VSp. As the number of investigators increases, given that they have different geographical, social and educational factors underlying their linguistic habits, we find that an increasing number of topological deformations is contributed, but because of their differences any conflation tends to a regular VSp. We assume that those criteria which are selected as being necessary by a given analyst, and not by another, will also be represented by transcriptional differences between them, and it is such data that we present in Table 1. Here, three analysts have transcribed stressed vowel nuclei in the same naturally read passage from four informants. Three of the informants were the

TABLE 1

Mean pair-disparity indices for three analysts and four informants (including the three analysts) by stressed V phonemes in continuous speech.

Informants

Phoneme	/ɪ/				/i/				/eɪ/				/ɛ/			
Pairs	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean
A	1	3	3	2·3	6	5	2	4·3	3	5	6	4·6	3	3	3	3·0
B	3	4	4	3·6	6	7	1	4·6	3	3	6	4·0	3	2	1	2·0
C	3	3	4	3·3	5	6	2	4·3	7	6	4	5·6	2	2	2	2·0
X	1	3	4	2·6	3	4	3	3·3	1	4	5	3·3	3	3	2	2·6
Means	2·00	3·25	3·75		5·00	5·50	2·00		3·50	4·50	5·25		2·75	2·50	2·00	

Phoneme	/æ/				/ɑ/				/ɒ/				/ɔ/			
Pairs	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean
A	1	2	2	1·6	1	3	2	2·0	2	2	2	2·0	4	3	1	2·6
B	3	1	3	2·3	1	1	0	0·6	2	3	2	2·3	5	3	3	3·6
C	4	1	4	3·0	0	1	1	0·6	2	3	2	2·3	4	3	2	3·0
X	3	2	2	2·3	0	2	2	1·3	2	2	2	2·0	1	2	1	1·3
Means	2·75	1·50	2·75		0·50	2·00	1·00		2·00	2·50	2·00		3·50	2·75	1·75	

Phoneme	/ʌ/				/ʊə/				/u/				/ɪu/			
Pairs	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean
A	1	2	1	1·3	6	6	6	6·0	4	5	1	3·3	4	1	3	2·6
B	3	2	2	2·3	12	11	14	12·3	3	2	4	3·0	3	3	3	3·0
C	6	7	3	5·3	9	7	11	9·0	4	4	0	2·6	4	2	2	2·6
X	3	1	3	2·3	4	3	3	3·3	4	5	3	4·0	2	2	4	2·6
Means	3·25	3·00	2·25		7·75	6·75	8·50		3·75	4·00	2·00		3·25	2·00	3·00	

Phoneme	/ɜ/				/aɪ/			
Pairs	*A-B*	*A-C*	*B-C*	Mean	*A-B*	*A-C*	*B-C*	Mean
A	3	2	3	2·6	3	5	4	4·0
B	3	3	3	3·0	6	4	2	4·0
C	2	2	4	2·6	3	2	3	2·6
X	4	2	3	3·0	5	5	2	4·0
Means	3·00	2·25	3·25		4·25	4·00	2·75	

analysts themselves. A comparison of the disparity between pairs of analysts on the basis of these data reveals the relative topological deformation of the hierarchically ordered subspace which contains the vowel profiles of individual informants in the central phase of the Survey. Table 1 presents the basic data. The disparity index for any pair of analysts x,y is

$$\sum_{1}^{m} \frac{(l_x - l_y)}{m}$$

where m is the number of realisations of the particular phoneme being examined, and l the diacritic locus given to any phoneme by a particular analyst. What counts as a relevant diacritic degree is determined empirically from the data, i.e., the range of a particular diacritic modification for a particular cardinal vowel in the transcriptions of a particular analyst. It is helpful to interpret this material, as we hinted above, by means of perspective differences. To do this we need to know the relative whole variety positions of analysts. These are given VERY crudely in terms of a linear scale of localisation in Figure 1. (We may characterise a non-localised variety as one

Figure 1 Linear representation of whole variety localisation of informants

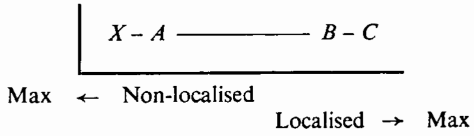

Fig. 1

which gives no speech-based indication of the area of origin of its speaker.) Our first expectation (Ringgaard 1965) is that the *A-B*, *A-C* disparities for informants *B*,*C*, should be relatively larger than the *B-C* ones. This is clearly true for /i/, /ɛ/, /ɔ/, /u/, /aɪ/, false for /ɪ/, /eɪ/ and uncertain for the rest. Similarly, we should expect the *A-B* disparity to be less than the *B-C* for informant *X*; this is true for /ɪ/, /eɪ/, /ɑ/, /ɪu/, equivocal for /ɔ/, /ɒ/, /ʌ/, and false for the rest. Without exhausting this material we must pass on to remark that there is a control which we should use to constrain the disparity indices. It may be that some disparity indices are artifacts arising from different reference norms or transcriptional conventions amongst the analysts. I.e., one analyst's [i] might not be the same as another's, either because the [i] may be articulated differently, or because of a wider or narrower diacritic range of usage. Therefore each analyst recorded a randomised series of cardinal vowels which were analysed by all analysts (including the performer). Disparity indices for these (single realisation values) are given in Table 2.

Again the disparities between analyst pairs on cardinal vowel and diacritic variability bear no simple LINEAR relationship to their whole variety differences. Table 3, which is a summary version of Table 1, is normalised by means of these cardinal vowel differences to give Table 4. Here again the patterning is complex and non-linear. Ringaard's 'sad conclusion' (1965) that the transcriptions of phoneticians

TABLE 2

Disparity indices for analysts, as performers and hearers, on random sequences of cardinal vowels[1]

	A-B	*B-A*	Pair mean[2]	*A-C*	*C-A*	Pair mean	*B-C*	*C-B*	Pair mean
i	1	0	1	0	1	1	1	2	1
e	0	1	1	1	5	3	3	4	3
ɛ	4	1	2	1	2	1	2	2	2
a	2	2	2	1	2	1	2	2	2
ɑ	1	2	1	1	1	1	4	0	2
ɔ	0	2	1	1	0	0	3	3	3
o	1	3	2	1	2	1	2	3	2
u	1	3	2	3	0	1	2	1	1
y	4	2	3	6	1	3	1	1	1
ø	3	3	3	2	6	4	2	5	3
œ	2	2	2	1	1	1	2	4	3
ɒ	5	1	3	3	3	3	2	3	2
ʌ	5	—	2	1	5	3	—	3	1
ɤ	1	5	3	1	8	4	3	7	5
ɯ	3	4	3	1	4	2	3	2	2
Means	2·20	2·21		1·60	2·73		2·28	2·80	
All CVs	2·20			2·16			2·55		
per pair	[CVs 1-8	1·50			1·37			2·25]	

[1] The performer appears first in each pair, the hearer second.
[2] Rounded down.

TABLE 3

Mean disparity index, by pairs, for four informants.

Informants	*A*	*B*	*C*	*X*
Pairs				
A-B	3·00	4·00	3·92	2·57
A-C	3·35	3·50	3·50	2·64
B-C	2·78	3·42	3·14	2·78

TABLE 4

Mean disparity index, by pairs, for four informants, normalised by cardinal vowel disparity scores (by pairs)[1]

Informants	*A*	*B*	*C*	*X*
Pairs				
A-B	1·50	2·50	2·42	1·07
A-C	1·98	2·13	2·13	1·27
B-C	0·53	1·17	0·89	0·53

[1] These values obtained from Table 2.

do not tell us so much about the speech of the locations in which they are working as about the speech of the phoneticians themselves, is only half the truth and not necessarily sad anyway.

The fact that speech perception, including, specifically, the derivation of extra-linguistic information, is multidimensionally complex, reflects the competence of the hearer as a TWO-way resonance system. Phonetic transcriptions CANNOT in this sense be true, or invariant; but if they are incorporated in research findings in a DYNAMIC fashion, as is suggested here, their topological deformations contribute further information which can only improve the ultimate inferences.

Department of English Language
University of Newcastle upon Tyne

REFERENCES

Ladefoged, P.
1960 "The Value of Phonetic Statements", *Language* 36:387.
Pellowe, J.N.H., G. Nixon, and V. McNeany
1971 "A Dynamic Modelling of Linguistic Variation: the Tyneside Linguistic Survey", to appear in *Lingua*.
Ringgaard, K.
1965 "The Phonemes of a Dialectal Area Perceived by Phoneticians and by the Speakers Themselves", in *Proceedings of the Vth International Congress of Phonetic Sciences*, Münster, 1964 (Basel, Karger).

DISCUSSION

BUSH (Stanford, Calif.)

May I ask if your analysts shared a common background in phonetic transcription training? If not, might that not account for at least as much potential variance as the factor of localized vs. non-localized? It would be important to specify the effect of training in this regard.

Is there any communality which has to do with similar training?

PELLOWE

We were not trained by the same means but even if we had been, I am sure that Ladefoged's (1960) "school effect" would not account for all the variability in these observations.

UNIDENTIFIED QUESTION

Is there any way of finding out what effect the analyst disparities have on the classification?

PELLOWE
We have in mind that one of the methods of determining this is for the three analysts to each analyse the same ten informants. These profiles are presented to the multi-dimensional variety space as if they were different speakers. The set or subset of optimal dynamic classification processes will then either place the three (same person) profiles closely or not; but in either case we shall have a measure of the local topological deformation due to analyst sociolinguistic differences and the measure will be attributable (by fractions) to individual criteria.

RICHTER (Bonn)
What will be the further prospects or continuation of your work?

PELLOWE
The survey is an open-ended monitoring system of linguistic change of any degree in phonological, prosodic (or 'suprasegmental') and grammatical systems. It is a general methodology applicable to any urban situation and not requiring native knowledge of either linguistic or social characteristics specific to (or diagnostic of) the locality. This is because the methodology is based upon iterative convergence techniques and analytic observation. Evaluation of an optimal classification of social-linguistic features is available both internally (mathematically) and externally (judgement tests). There are many extensions both of data collection and of classificatory technology which we have in mind as relevant to longitudinal (through time) and latitudinal (through space) aspects of language variability surveys.

RICHTER
Will tolerances, which have to be admitted in your opinion, be determinable? Could you imagine something like a flexible normation of phonetic transcription?

PELLOWE
The tolerances of analytic variability amongst phoneticians are determinable, IN THE CONTEXT OF A SPECIFIED RANGE OF DATA DIFFERENCES, by an extension of the methods sketched in this paper. That is, if all analysts code the same interview from the same informant and we input these different analyses AS IF they were DIFFERENT informants, the different classified positions of these analyses will be a defined index of relevant analyst differences, for that particular part of the multidimensional subspace. I think that it is obvious that analyst differences will NOT be invariant in different subspaces of the multidimensional variety space, and your phrase 'flexible normation' captures this quite well.