THE DIMENSIONALITY OF GRAMMATICAL VARIATION

DAVID SANKOFF *Université de Montréal* and HENRIETTA J. CEDERGREN *Université du Québec à Montréal*

Recent approaches to variability have focused on the rank ordering of linguistic objects (e.g. noun-phrase transformations, environments of phonological reduction rules) or sociolinguistic objects (e.g. speakers, speech communities). These one-dimensional orderings can take various forms, such as implicational scales, squishes, or the ranking of variable constraints. It is suggested here that sometimes a higher dimensional configuration might better represent the variation. Computer-based heuristic procedures, notably multidimensional scaling techniques, are available for investigating the dimensionality of large data sets. In applying such methods to some well-known data, we confirm the one-dimensionality of Ross' collection of tests of noun-phrasiness as well as of the classes of fake NP's; and we confirm the two-dimensionality of the complementizer QUE-deletion rule in Montreal French. The linguistic environment of syllable-final S-reduction for speakers of Panamanian Spanish is shown to be three-dimensional. Phonological data are used to find two sociolinguistically significant dimensions of the speech community.*

To cope with regular variation in grammar, and within and between speech communities, generative grammar has been extended through the introduction of a number of analytical devices, including variable rules (Labov 1969; Cedergren & Sankoff 1974), implicational scales (Elliott, Legum & Thompson 1969; DeCamp 1971; Bickerton 1971), and squishes (Ross 1972, 1973; Sag 1973). Much of this work has been concerned with the ordering, or ranking, of linguistic entities according to some criterion.

For example, Labov ordered various categories of copula complements according to the extent to which they favor copula contraction and deletion. Elliott, Legum & Thompson ordered a number of sentences with subject pronouns deleted from *while*-clauses, according to their acceptability to informants. Ross ordered 'fake' noun phrases such as expletive *there* and weather *it* according to the applicability of a battery of tests of 'noun-phrasiness'. In all these cases, and in most others which have been published, the ordering is essentially two-dimensional. Thus, in the course of ordering fake NP's, Ross simultaneously constructed an order for a number of rules applicable to NP's—such as tag formation, conjunction reduction, *being* deletion etc.—according to their breadth of application. Elliott, Legum & Thompson's acceptability ordering is concurrent with an ordering of informants, according to the strictness of their judgments of acceptability. And Labov, while ordering copula complements, simultaneously ranked phonological environments of the copula according to their effects on contraction and deletion. Despite the prevalence of two-dimensional examples, however, there is no reason to exclude one-, three-, or higher-dimensional orderings from the analysis and understanding of linguistic variation. The formalism, methodology, and interpretation of multi-dimensionality are the objects of this paper.

* An earlier version of this paper appeared as Technical Report 438 of the Centre de Recherches Mathématiques, Université de Montréal. We are grateful to J. B. Kruskal for comments on that earlier version.

1. THE CASE OF TWO DIMENSIONS. Criteria such as relative frequency of usage, rule application probability, acceptability, and grammaticality can be represented numerically, either by a scale from zero to one, or simply by the two values zero and one, or by some intermediate type of scale. Suppose that such a criterion x varies as in Table 1 depending on which of a set of categories, contexts, or informants—represented by i varying between 1 and m—is applicable. Suppose further that x

INDEX i \ INDEX j	1	2	3		j_2	j_1	n
1	$x(1,1)$	$x(1,2)$	$x(1,3)$	$\cdots$			
2	$x(2,1)$	$x(2,2)$	$\cdots$				
3	$x(3,1)$	$\vdots$					
$\vdots$	$\vdots$						
i_2					$x(i_2, j_2)$		
i_1						$x(i_1, j_1)$	
m							

TABLE 1. Scaling condition.

i INDEX \ j INDEX	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2
3	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	3	3	3	4
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	3	3	3	5
5	1	1	2	1	2	1	1	2	2	2	1	3	3	3	3	3	4	4	5
6	1	1	2	2	3	1	2	2	4	4	2	4	4	5	5	4	5	5	5
7	1	–	2	5	2	2	3	3	–	2	5	–	–	5	5	5	4	5	5
8	1	–	2	2	1	5	3	3	–	5	5	–	–	5	5	5	3	3	5
9	1	–	3	5	3	3	4	4	–	3	5	–	–	5	5	2	5	5	5
10	1	–	2	4	3	5	1	5	–	5	5	–	–	5	5	5	5	4	5
11	2	4	3	3	4	5	4	6	5	5	5	4	5	5	5	5	5	5	5
12	4	2	4	3	5	4	4	5	5	5	5	5	5	5	5	5	5	5	5

TABLE 2. Ross' Fake NP Squish.
O.K. = 1; ? = 2; ?? = 3; ?* = 4; * = 5; ** = 6; – = does not apply.

i INDEX
1 Animates
2 Forces of nature; concretes
3 Events
4 Abstracts
5 *Tack*
6 *Headway*
7 *It* (*be muggy*)
8 *It* (S)
9 *It* (rain)
10 *There*
11 *Tabs*
12 *Heed*

j INDEX
1 Tag formation
2 Head of relative clause
3 Inside derived nominals
4 *Get* passives
5 *To be* deletion
6 Conjunction reduction
7 Acc-*ing*
8 ...'*s*
9 Pronominalization
10 Equi
11 Tough mvt., topicalization & swooping
12 NP shift
13 Right node raising
14 Left and right dislocation
15 *Being* deletion
16 *What's ... doing X*
17 *Think of ... as X*
18 Double raising
19 Promotion and subj. of *be prevented*

also depends on which of a second, independent set of indices—j varying between 1 and n—is present. Then we say that the two sets of indices are scaled if the two-dimensional scaling condition holds, i.e. if $x(i_1, j_1)$ is not less than $x(i_2, j_2)$ whenever i_1 and j_1 are not less than i_2 and j_2, respectively.

For example, i might index the twelve classes of nouns, NP's, and fake NP's which Ross 1973 ordered—from concrete animate nouns at one extreme, to *heed*, *tabs*, and expletive *there* at the other—while j indexes his nineteen tests of noun-phrasiness, from the most general to the most rigorous. Then suppose, as in Table 2, that $x(i, j) = 1$ represents the fact that class i passes test j, and $x(i, j) = 6$ means that the application of test j to class i leads to a doubly starred ungrammatical result, while values 2 to 5 represent decreasing degrees of grammaticality. Then, to the extent that the two-dimensional scaling condition holds for i_1 running from 1 to 12, i_2 running from 1 to i_1, j_1 running from 1 to 19, and j_2 running from 1 to j_1, we may say that both the tests and the fake NP classes are scaled.

Given some criterion x, and its dependence on two intersecting sets of categories, contexts etc., there are standard statistical techniques, e.g. Guttman's scalogram analysis (Torgerson 1958, ch. 12), for finding an indexing of the two sets which satisfies the scaling condition, or satisfies it as much as possible. Thus, were the Fake NP Squish data presented as in Table 3, it would not be difficult to find

	Acc-*ing*	...'s	Promotion and subj. of *be prevented*	Head of relative clauses	*Being* deletion	*Get* passives	Tough mvt., topicalization and swooping	*What's ... doing X*	*To be* deletion	Equi	*Think of ... as X*	Pronominalization	Tag formation	Double raising	NP shift	Left and right dislocation	Conjunction reduction	Right node raising	Inside derived nominal
It (rain)	4	4	5	–	5	5	5	2	3	3	5	–	1	5	–	5	3	–	3
Forces of nature; concretes	1	1	2	1	1	1	1	2	1	1	2	1	1	2	1	1	1	1	1
Heed	4	5	5	2	5	3	5	5	5	5	5	5	4	5	5	5	4	5	4
It (*be muggy*)	3	3	5	–	5	5	5	5	2	2	4	–	1	5	–	5	2	–	2
Headway	2	2	5	1	5	2	2	4	3.	4	5	4	1	5	4	5	1	4	2
Abstracts	1	1	5	1	2	1	1	3	1	1	3	1	1	3	1	1	1	1	1
Animates	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
It (S)	3	3	5	–	5	2	5	5	1	5	3	–	1	3	–	5	5	–	2
Events	1	1	4	1	1	1	1	3	1	2	3	1	1	3	1	1	1	1	1
Tack	1	2	5	1	3	1	1	3	2	2	4	2	1	4	3	3	1	3	2
Tabs	4	0	5	4	5	3	5	5	4	5	5	5	2	5	4	5	5	5	3
There	1	5	5	–	5	4	5	5	3	5	5	–	1	4	–	5	5	–	2

TABLE 3. Ross' Fake NP Squish before appropriate indexing.

permutations of the rows and columns which would result in the scale of Table 2, or something very similar.

Before this problem even arises, however, a more basic question must be answered: given that x depends on a set of entities or situations (e.g. possible transformations of sentences, copula environments, acceptability judgments), does a linguistically significant pair of dimensions exist (e.g. NP's and fake NP's vs. tests, complement categories vs. phonological environments, informants vs. sentences), on which we can cross-classify this set? Thus we have two rather incommensurable ways of assessing a two-dimensional scale: linguistic significance, and the degree to which the scaling condition is satisfied.

2. The general case. Suppose that x depends on d sets of indices $i, j, \ldots, z$— where index i, say, runs from 1 to n_i, and $x(i_1, j_1, \ldots, z_1)$ is not less than $x(i_2, j_2, \ldots, z_2)$ whenever $i_1, j_1, \ldots, z_1$ are not less than $i_2, j_2, \ldots, z_2$, respectively. This is illustrated in Figure 1.

0	0	0	0	1	1	1	1	1	1

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1
0	0	0	0	0	0	0	1	1	1
0	0	0	0	1	1	1	1	1	1
0	0	1	1	1	1	1	1	1	1
0	0	1	1	1	1	1	1	1	1
0	0	1	1	1	1	1	1	1	1
0	0	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1

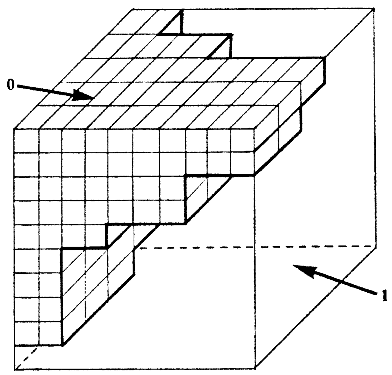


Figure 1. From top to bottom: one-, two-, and three-dimensional scales for a variable x taking on values zero and one.

Then we say that the d-dimensional scaling condition is satisfied. Once again, given a way to calculate or measure x in a large number of situations, the major problem is this: is there some number d, and some meaningful d-dimensional classification of these situations, such that the d-dimensional scaling condition is satisfied or nearly satisfied?

In the next sections we investigate this question for three sets of linguistic data. Before doing this we sketch the general procedure, which is also schematized in Figure 2.

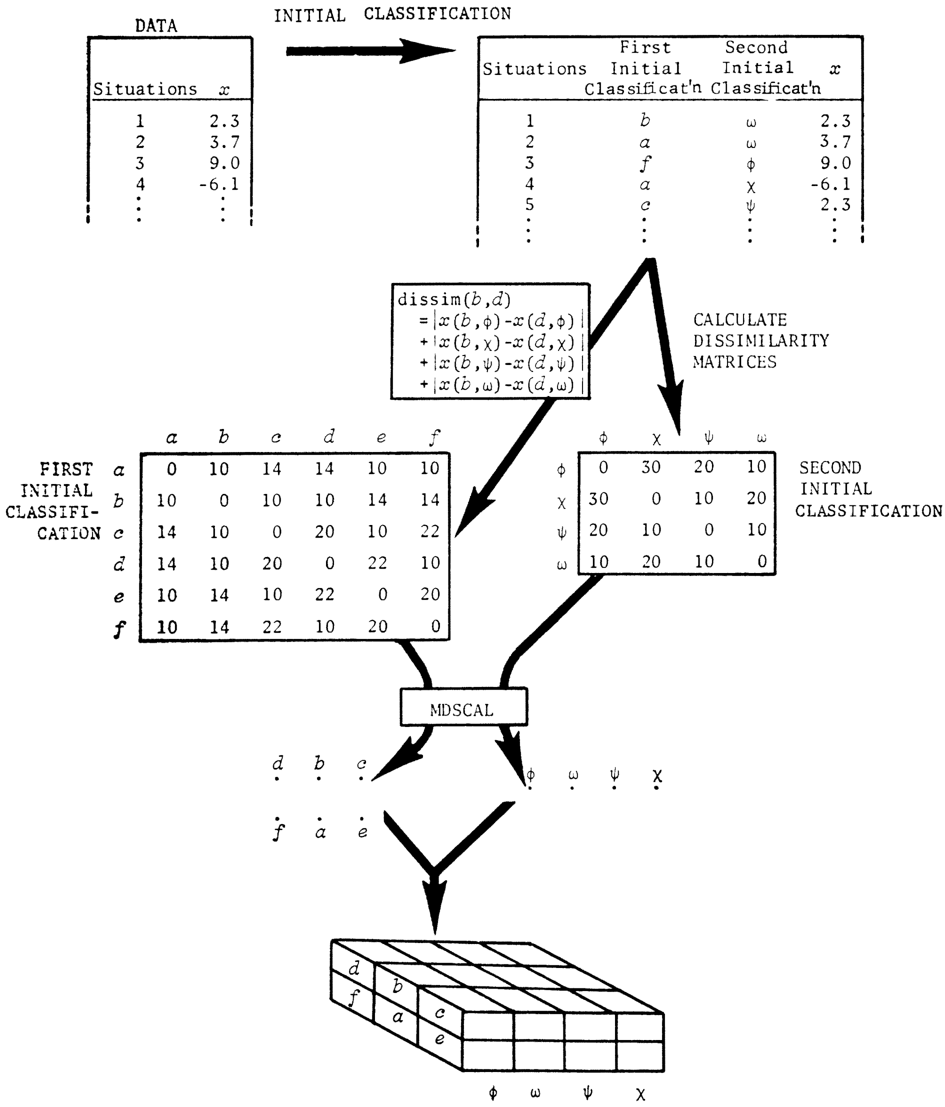


FIGURE 2. Schema of procedure for producing multidimensional scales.

Suppose we have measured x in a number of situations. The method depends on the existence of an initial two-way cross-cutting classification (not necessarily a scale) of the situations, e.g. informants vs. linguistic contexts, or fake NP's vs.

tests as in Table 3. As with these examples, this initial two-way classification of the situation usually arises naturally in the problem being studied. The next step is to find the appropriate dimensionality for the set of classes or categories *within* one or both of the two initial ways of classifying the situations. If this results in a d_1-dimensional structure for the categories in the first way and a d_2-dimensional array for those in the other, we might then hope to construct a d_1+d_2-dimensional scale for the entire data set. In practice, however, the major interest is in structuring one or both of the initial sets of categories.

The problem of finding the dimensionality of the rows, say, or of the columns of a two-way array such as Table 3, is well-known in statistics. It can be approached in many different ways—such as factor analysis, principal components analysis, unfolding, multidimensional scaling, and related procedures for decomposing, reducing, and representing data matrices. (For a bibliography and a comparison of these methods, see Shepard 1972a.) In this paper we will take an approach using multidimensional scaling, which is applicable to many different types of data, and which has three basic steps. Consider either one of the two initial ways of classifying the situations. First, using the given values of x, we calculate a dissimilarity matrix measuring how different the categories are within this classification. Second, we construct scales in 1, 2, 3, and higher dimensions most consistent with the dissimilarity matrix. Finally, using statistical and linguistic criteria, we select the best dimensionality and associated scale.

Two statistical constructions have been mentioned, for the dissimilarity matrix and for the d-dimensional scales. The dissimilarity matrix contains entries *dissim* (i_1, i_2) which measure how different category i_1 and category i_2 are in one of the initial classifications. Basically, this is calculated by summing the differences between the values taken by x when indexed by i_1 and by i_2, over all possible categories in the other initial classification. This sum may be symbolized as:

$$|x(i_1, 1) - x(i_2, 1)| + |x(i_1, 2) - x(i_2, 2)| + |x(i_1, 3) - x(i_2, 3)| + \ldots$$

The reason we need two cross-cutting initial classifications in the first place is specifically to be able to form this sum. In different examples, the sum may be

FAKE NP's	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0.21	0.53	0.58	1.32	2.21	2.60	2.53	2.87	3.00	3.47	3.53
2	0.21	0	0.32	0.37	1.11	2.00	2.33	2.27	2.60	2.73	3.26	3.32
3	0.53	0.32	0	0.16	0.79	1.68	1.93	1.87	2.33	2.33	2.95	3.00
4	0.58	0.37	0.16	0	0.74	1.63	1.87	1.80	2.27	2.27	2.90	2.95
5	1.32	1.11	0.79	0.74	0	0.90	1.27	1.60	1.67	1.67	2.16	2.21
6	2.21	2.00	1.68	1.63	0.90	0	0.93	1.13	1.07	1.07	1.26	1.32
7	2.60	2.33	1.93	1.87	1.27	0.93	0	0.87	0.67	0.93	1.13	1.27
8	2.53	2.27	1.87	1.80	1.60	1.13	0.87	0	1.27	0.73	0.93	1.20
9	2.87	2.60	2.33	2.27	1.67	1.07	0.67	1.27	0	0.93	0.87	1.00
10	3.00	2.73	2.33	2.27	1.67	1.07	0.93	0.73	0.93	0	0.60	0.87
11	3.47	3.26	2.95	2.90	2.16	1.26	1.13	0.93	0.87	0.60	0	0.47
12	3.53	3.32	3.00	2.95	2.21	1.32	1.27	1.20	1.00	0.87	0.47	0

TABLE 4. Matrix of dissimilarities between fake NP's.

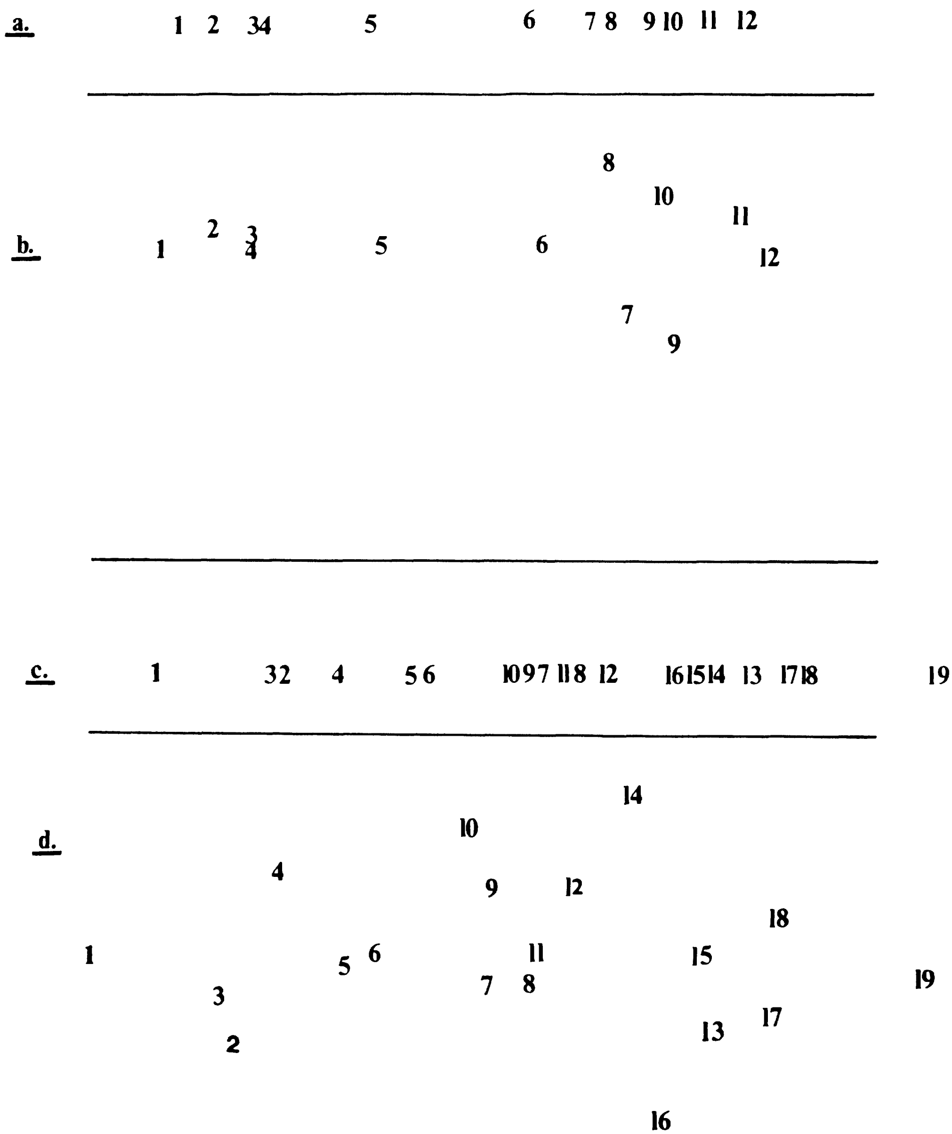


FIGURE 3. (a) One-dimensional scale of fake NP's; (b) Two-dimensional scale of fake NP's; (c) One-dimensional scale of NP tests; (d) Two-dimensional scale of NP tests.

weighted, normalized, or otherwise slightly altered to take into account the quantity and quality of the data involved.

The second statistical technique is the construction of a *d*-dimensional scale based on the dissimilarity matrix. This is a much more involved procedure, and to carry it out we use MDSCAL, a multidimensional scaling computer program (Kruskal 1964a,b) which has been widely tested in diverse applications. In linguistics, techniques of multidimensional scaling have been used in dialectology (Black 1973) and very extensively in acoustic phonetics (Terbeek 1973, Shepard 1972b). The MDSCAL program uses a method of successive approximation to construct a *d*-dimensional configuration of the categories in which those categories which are furthest apart (as measured by *d*-dimensional Euclidean distance) are, insofar as is possible, those which have the greatest dissimilarity coefficients, while those which are closest together have the smallest dissimilarities.

3. AN EXAMPLE. To illustrate the method, we first use a data set for which the outcome can be fairly well anticipated, namely Ross' Fake NP Squish data, as shown in Table 2. The initial classification is fake NP classes vs. tests, and we would expect our method to produce simply a one-dimensional order in both these dimensions. In addition, since the input to the method consists only of the values of x as in Table 3, and not the ordering of the categories, we would want it to output orderings which correspond as much as possible to those of Ross.

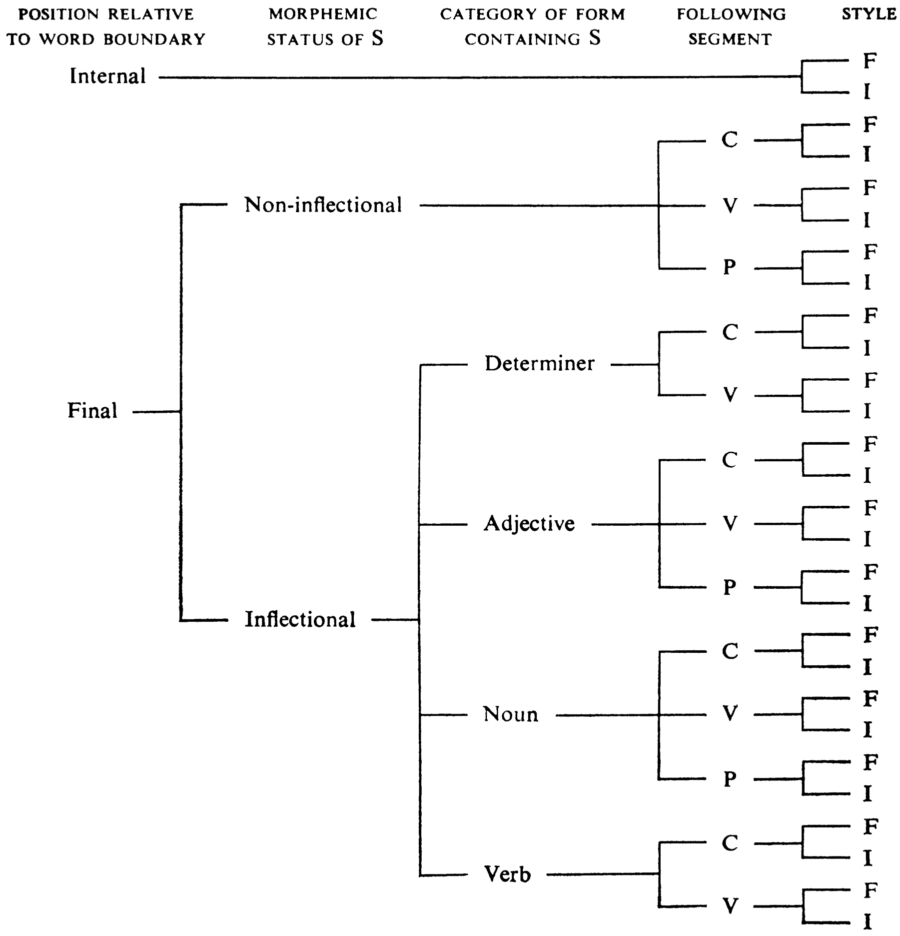


TABLE 5. Classification of 28 environments of S.
C, V, P = consonant, vowel, pause; F, I = formal, informal.

First, let us examine the twelve classes of fake NP's. For each pair of classes i_1 and i_2, we sum $|x(i_1, k) - x(i_2, k)|$ over all tests k for which values are given; and to take into account that some values are missing, we divide by the number of terms in the sum. This results in the dissimilarity matrix in Table 4 (p. 168).

Fed into MDSCAL, this matrix produces the one-dimensional scale in Figure 3a (p. 169); thus the method produces exactly the same ranking as Ross assigned.

However, MDSCAL can also produce two-, three-, or higher-dimensional configurations of these classes. How can we know that these contain no further interesting information about the fake NP-classes? Let us look at the two-dimensional scale in Figure 3b. It is clear that, despite our giving the computer two dimensions in which to construct a scale, it essentially used only one, the second being largely superfluous.

Next we examine the results of similarly analysing the second initial dimension, containing the different tests of noun-phrasiness. Once again, Figure 3c shows that the one-dimensional scale output by MDSCAL is consistent with Ross', though with a few differences; and Figure 3d shows that by far the largest proportion of variation in a two-dimensional scale is contained along one dimension. Thus, since the fake NP's and the tests can each be ordered along one dimension, our procedure gives a two-dimensional scale very similar to Table 2 as proposed by Ross.

4. Some higher-dimensional examples. Cedergren 1973a counted occurrences of three variants of the Panamanian Spanish syllable-final sibilant S in the 28 stylistically and/or grammatically different environments listed in Table 5, for each of 79 speakers. The first initial classification is according to the environments and the second, by speakers.

For each environment we calculated X, Y, and Z, the proportion of each of the three variants, for each speaker. Then for each pair of environments i_1 and i_2, the dissimilarity was calculated in terms of the difference between their X values, the difference between their Y values, and the difference between their Z values, suitably weighted to take into account the total number of occurrences in the environment, and combined over all speakers. (See the Appendix for the weighting procedure.)

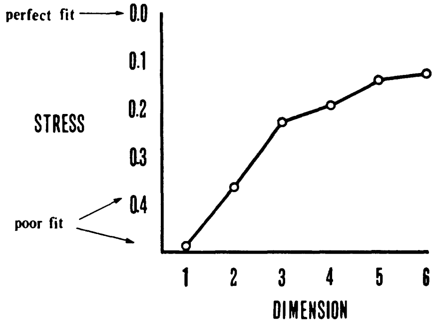
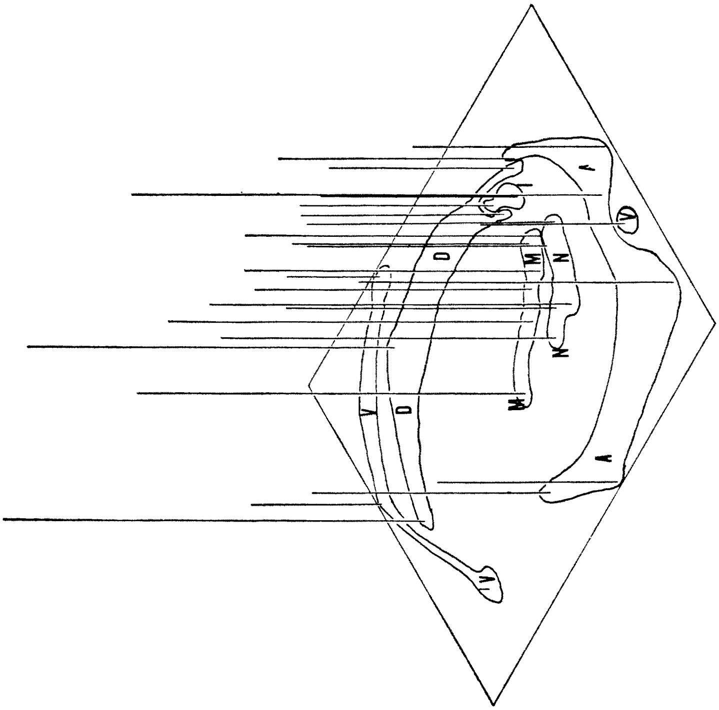
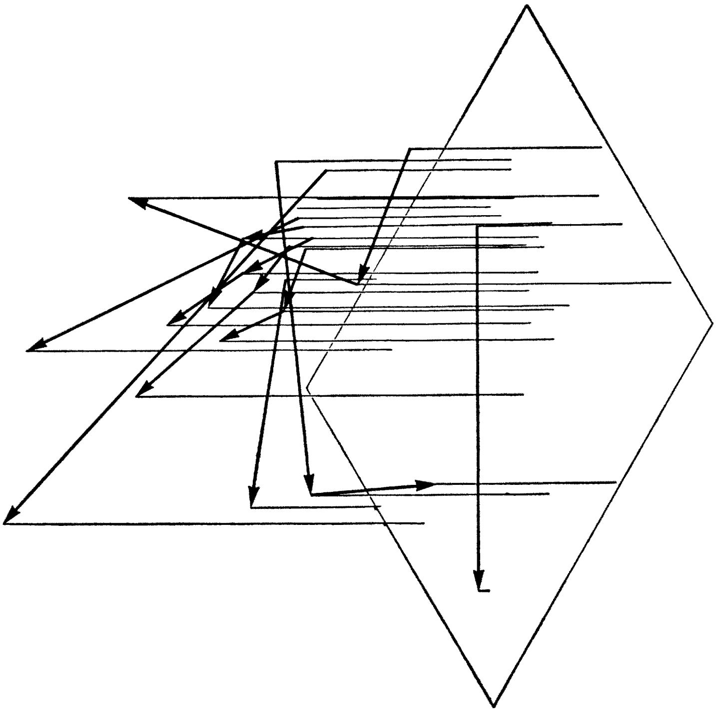


Figure 4. Stress curve for scales of S environments showing 'elbow' at dimension = 3.





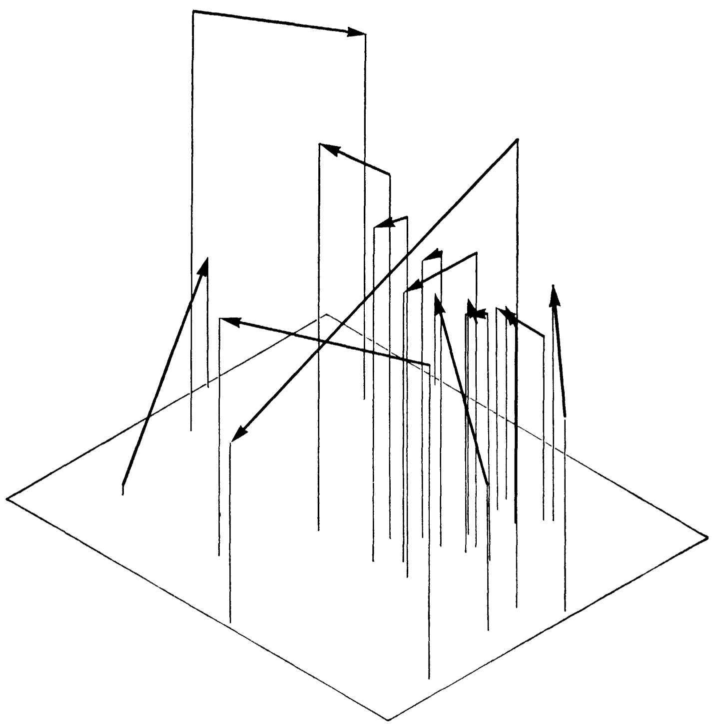


FIGURE 5. Three-dimensional scale of S environments. Upper tips of vertical lines represent position of environments in space.

Top left: Single arrows go from consonantal environments to corresponding vowel environments. Two arrows in sequence proceed from consonant to vowel to pause. S-reduction increases from upper left to lower right of diagram.

Bottom left: Arrows go from informal to corresponding formal context. S-reduction tends to increase from left to right.

Top right: Vertical lines corresponding to environments grouped according to position of S in word and category of form containing S. V = verb, D = determiner, I = internal, M = non-inflected, N = noun, A = adjective. S-reduction increases from back to front of diagram.

The output from MDSCAL was then obtained for one to six dimensions. One useful feature of MDSCAL is that it measures, in terms of a quantity called 'stress', how well the scale fits the dissimilarity data input in each dimension. As the dimension increases, it is always possible to make a scale which fits better than a lower-dimensional scale; so unless the increase in fit is dramatic, we consider that increasing the dimensionality does not provide us with a better model. In this case, as is clear in Figure 4, (p. 171) the increase in fit is quite marked as we go from one to two dimensions, and from two to three dimensions, but there is little improvement from then on.

Thus we choose a three-dimensional scale for our environments, as shown in Figure 5 (pp. 172–3). Here we see that one dimension accounts for variation according to the following phonological segment, while a second dimension accounts for variation according to morphological status. The third dimension seems to capture stylistic variation. The position of the various environments in this three-dimensional configuration is quite consistent with variable-rule analyses of these same data (Cedergren 1973a,b).

A two-dimensional projection of the three-dimensional configuration of Fig. 5 is shown in Figure 6: in Fig. 6a the environments have been labeled according to the following phonological segment to show the scaling from consonant to vowel to pause, while Fig. 6b is the same projection labeled according to morphological status.

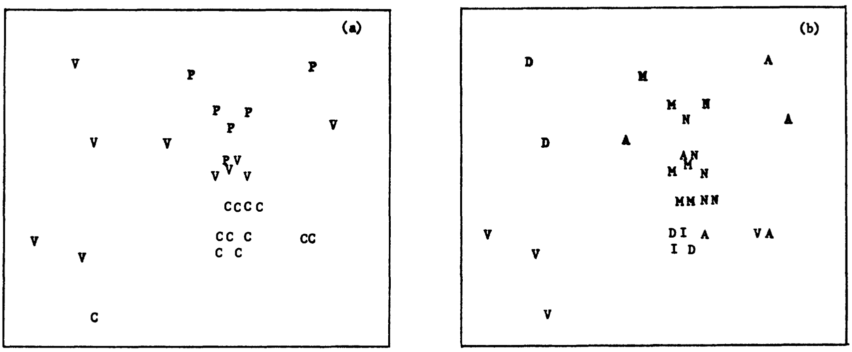


FIGURE 6. Two-dimensional projection of configuration of Figure 5.

(a) Labeled according to following phonological segment (and showing ordering from bottom to top) of consonant, vowel, and pause environments.

(b) Labeled according to morphological status—showing verbs and determiners mostly at the extreme left, adjectives mostly at the extreme right, and other forms in the center.

This shows a general tendency to order from verbs and determiners, at one extreme, to adjectives at the other—with noun forms, non-inflectional word-final environments, and word-internal environments clustering in the center. The third dimension of variation, style, cannot of course be seen in the same two-dimensional projection as the other two dimensions; but other projections (not shown) of the configuration in Fig. 5 do portray the systematic distinction between formal and informal environments.

We did not attempt to restructure the second initial classification, i.e. the speakers, by comparing pairs of speakers in all the environments of the single variable S. This would have led predictably to a one-dimensional scale, since any speaker's distribution of the three S variants in the different environments is more or less determined by his over-all rate of S-reduction. Instead, we compared speakers' behavior not only with respect to S-reduction, but also R-reduction, N-velarization, CH-lenition, and D-deletion simultaneously. (These additional data are also dis-

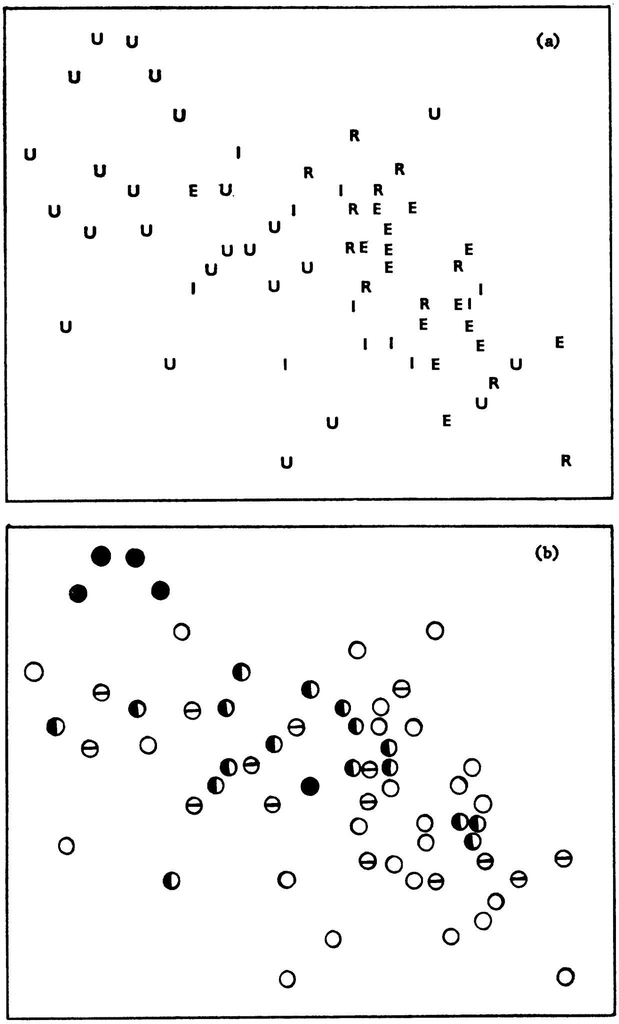


FIGURE 7. Two-dimensional scale of 62 Panamanian Spanish speakers.

(a) Speakers labeled according to age of arrival in city. U = city-born, I = arrived during childhood, E = adolescent arrivals, R = adult arrivals.

(b) Speakers labeled according to socio-economic level. Filled-in circles, half-filled circles, striped circles, and open circles represent upper, upper-middle, lower-middle, and lower levels, respectively.

cussed by Cedergren 1973a.) We used only 62 of the 79 speakers because of computer storage limitations.

In this case, the stress curve has no 'elbow' which would help us fix the dimensionality. However, we examined the two-dimensional classification of speakers, and found that we were able to interpret one of the dimensions as ordering the speakers according to their age at arrival in Panama City from the rural areas as in Figure 7a (p. 175), and the second dimension as ordering them according to social class, as in Figure 7b.

Note that this demographic information was not given to the computer; nonetheless, it was able to arrange the speakers in this reasonable way solely on the basis of their phonological behavior.

The results in Figs. 5 and 6 prompt this question: to what extent can the 79 speakers, taken individually, be considered to structure the space of environments in the same way? We do not investigate this here, but methods of multidimensional scaling are available which would take individual differences into account (Carroll & Chang 1970).

5. *Que*-DELETION. A much-analysed set of data is that on complement *que*-deletion in Montreal French (G. Sankoff 1974), consisting of frequencies of deletion and non-deletion of *que* in nine phonological environments for sixteen speakers. Cedergren & Sankoff 1974 showed, on the basis of *que*-deletion, how the phono-

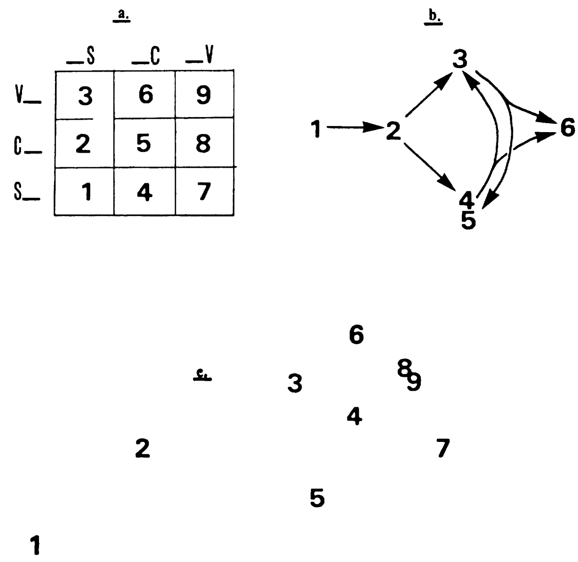


FIGURE 8

(a) Two-dimensional arrangement of environments of *que* in variable-rule model.

(b) Order in which *que*-deletion spreads from environment to environment, abstracted from Bickerton's analysis (1973).

(c) Two-dimensional scale produced by multidimensional scaling.

logical environments could be classified in two dimensions: preceding vs. following segment, both ranked as vowels, non-sibilant consonants, and sibilants as in Figure 8a. Bickerton 1973 argued that a more natural categorization of the environments would be as in Figure 8b.

We calculated deletion proportions in each environment, as well as a dissimilarity matrix for the environments much as in the Fake NP Squish example. Figure 8c shows the two-dimensional MDSCAL output. Of Cedergren & Sankoff's two dimensions, the succeeding phonological environment is clearly confirmed, and the preceding environment also emerges, although not as clearly. Interestingly enough, the pattern that emerges is also thoroughly consistent with Bickerton's schema.

6. CONCLUSION. Multidimensional scaling is a valuable heuristic tool for investigating the structure of data on grammatical variability. Its advantage is that it makes few assumptions about the structures which generated the data, such as are required for variable-rule analysis or more familiar scaling procedures. It is dependent, of course, on the analytic grid used in collecting the data; but if this is fine enough, little bias is introduced.

The disadvantages lie in the fact that the output tends to be less clear-cut and interpretable than analyses according to more structured models. However, when one has large data sets and some experience with the procedure, this is a preferable method for finding dimensionality, prior to any more structured analysis.

APPENDIX: DISSIMILARITIES AND WEIGHTS

Let $s(i, k)$, $h(i, k)$, and $\phi(i, k)$ be the number of occurrences of the three variants of S, and $n(i, k) = s(i, k) + h(i, k) + \phi(i, k)$ for speaker i in environment k. Then to calculate the dissimilarity between two environments k_1 and k_2, we first calculate a weighting factor $w_i(k_1, k_2) = \sqrt{n(i, k_1)n(i, k_2)}$ and then sum up

$$w_i(k_1, k_2)\left[\left(\frac{s(i, k_1)}{n(i, k_1)} - \frac{s(i, k_2)}{n(i, k_2)}\right)^2 + \left(\frac{h(i, k_1)}{n(i, k_1)} - \frac{h(i, k_2)}{n(i, k_2)}\right)^2 + \left(\frac{\phi(i, k_1)}{n(i, k_1)} - \frac{\phi(i, k_2)}{n(i, k_2)}\right)^2\right]$$

over all speakers i. This sum is then normalized by dividing by $w_1(k_1, k_2) + w_2(k_1, k_2) + \cdots + w_{79}(k_1, k_2)$. The square root of the resultant quantity is taken to be *dissim*(k_1, k_2).

To calculate dissimilarities between two of the 62 speakers i_1 and i_2, over all environments for the five phonological variables, we first compute the weight $w_k^S(i_1, i_2) = \sqrt{n(i_1, k)n(i_2, k)}$ for each environment k for the S data, and similar quantities $w_k^R(i_1, i_2)$ etc. for the other variables. Then we sum up

$$w_k^S(i_1, i_2)\left[\left(\frac{s(i_1, k)}{n(i_1, k)} - \frac{s(i_2, k)}{n(i_2, k)}\right)^2 + \left(\frac{h(i_1, k)}{n(i_1, k)} - \frac{h(i_2, k)}{n(i_2, k)}\right)^2 + \left(\frac{\phi(i_1, k)}{n(i_1, k)} - \frac{\phi(i_2, k)}{n(i_2, k)}\right)^2\right]$$

over all k plus analogous sums for the R and other data. The total of $w^S + w^R + \cdots$ is used for normalizing, and *dissim*(i_1, i_2) is given by the square root of the normalized sum.

REFERENCES

BAILEY, C.-J. N., and R. W. SHUY (eds.) 1973. New ways of analyzing variation in English. Washington, DC: Georgetown University Press.

BICKERTON, DEREK. 1971. Inherent variability and variable rules. Foundations of Language 7.457–92.

——. 1973. Quantitative versus dynamic paradigms: the case of Montreal QUE. In Bailey & Shuy, 23–43.

BLACK, PAUL. 1973. Multidimensional scaling applied to linguistic relationships. Murray Hill, NJ: Bell Laboratories.

CARROLL, J. D., and J.-J. CHANG. 1970. Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckart-Young decomposition. Psychometrika 35.283–319.

CEDERGREN, HENRIETTA. 1973a. Interplay of social and linguistic factors in Panama. Cornell dissertation.

——. 1973b. On the nature of variable constraints. In Bailey & Shuy, 13–22.

——, and DAVID SANKOFF. 1974. Variable rules: performance as a statistical reflection of competence. Lg. 50.333–55.

DECAMP, DAVID. 1971. Toward a generative analysis of a post-creole speech continuum. Pidginization and creolization of languages, ed. by D. Hymes, 349–70. Cambridge: University Press.

ELLIOTT, DALE; STANLEY LEGUM; and SANDRA A. THOMPSON. 1969. Syntactic variation as linguistic data. Papers from the 5th Regional Meeting, Chicago Linguistic Society, 52–9.

KRUSKAL, J. B. 1964a. Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis. Psychometrika 29.1–27.

——. 1964b. Nonmetric multidimensional scaling: a numerical method. Psychometrika 29.115–29.

LABOV, WILLIAM. 1969. Contraction, deletion, and inherent variability of the English copula. Lg. 45.715–62.

ROSS, JOHN R. 1972. The category squish: Endstation Hauptwort. Papers from the 8th Regional Meeting, Chicago Linguistic Society, 316–28.

——. 1973. A fake NP squish. In Bailey & Shuy, 96–140.

SAG, IVAN. 1973. On the state of progress on progressives and statives. In Bailey & Shuy, 83–95.

SANKOFF, GILLIAN. 1974. A quantitative paradigm for the study of communicative competence. Explorations in the ethnography of speaking, ed. by R. Bauman & J. Sherzer, 18–49. Cambridge: University Press.

SHEPARD, R. N. 1972a. A taxonomy of some principal types of data and of multidimensional methods for their analysis. Multidimensional scaling: theory and applications in the behavioral sciences, ed. by R. N. Shepard, A. K. Romney & S. B. Nerlove, 21–47. New York: Seminar Press.

——. 1972b. Psychological representation of speech sounds. Human communication: a unified view, ed. by E. E. David & P. B. Denes, 67–113. New York: McGraw-Hill.

TERBEEK, DALE. 1973. Six dimensions of vowel quality. Papers from the 9th Regional Meeting, Chicago Linguistic Society, 672–8.

TORGERSON, W. S. 1958. Theory and methods of scaling. New York: Wiley.

[Received 11 July 1975.]