

## PERCEPTUAL ANALYSIS FROM CONFUSIONS BETWEEN VOWELS

L. J. TH. VAN DER KAMP

*Psychological Institute, University of Leyden, The Netherlands*

and

L. C. W. POLS

*Institute for Perception, Soesterberg, The Netherlands*

### ABSTRACT

In an experiment on vowel identification confusions were obtained between 11 Dutch vowel sounds. To recover the perceptual configurations of the stimuli multidimensional scaling techniques were applied directly to the asymmetric confusion matrix, and to the symmetrized confusion matrix. In order to investigate the agreement between the different solutions, all the configurations were matched to a perceptual configuration obtained in a previous study on the perception of the same vowel sounds in which similarities were obtained by the method of triadic combinations. The different solutions turned out to be very similar and in close agreement with the perceptual structure from similarity judgements. The solutions obtained by discarding the asymmetries first, showed the best correspondence with the perceptual similarity structure. In particular the solution based on Luce's choice axiom to correct for response bias seemed to be the most attractive.

### 1. SIMILARITY ANALYSIS OF VOWEL SOUNDS

In a previous study on the perception of Dutch vowel sounds (POLS et al., 1969) multidimensional scaling of similarity judgements resulted in the finding that vowels can be mapped in a 3-dimensional perceptual space. The similarities were obtained from a group of subjects by the method of triadic combinations (TORGERSON, 1958). This means that the stimuli were presented to the subject in groups of three, and the subject decides which of the three possible pairs of stimuli are most similar and which are most different.

This is one way of defining similarity between stimuli. Another way to define similarity is in terms of confusability: two stimuli then are said to be similar to the extent that they tend to evoke the same behavior, that is that they tend to be confused (MILLER, 1956). This approach in the

study of the perception of speech is perhaps more realistic than a direct judgement of similarities: more realistic in the sense that instead of relative comparisons of possible combinations of stimuli the observer is now presented with a single stimulus which has to be identified.

The main reason for supplementing our previous study is to investigate whether the psychological space inferred from triadic similarity judgements has the same underlying structure as from identifications. Our procedure will be that we take the similarity space of our previous study as the target, and find out to what extent an identification space can be matched with this target.

In addition, attention is given to some methodological problems. These problems have to do with the fact that the initial confusion matrix is asymmetric. A number of current techniques for multidimensional analysis assume a symmetric matrix of similarities. Is it justified to symmetrize the confusion matrix beforehand, and, if so, by what method? In this study three methods will be used for symmetrizing the confusion matrix; one also corrects for response bias, the other two do not.

## 2. EXPERIMENT

As stimuli 11 Dutch vowel-like sounds were used. The signals were generated by a digital computer (PDP-7) and made audible by a digital-to-analog converter. In fact, the signals are continually generated single periods. These periods were taken from the constant vowel parts of normally spoken *h* (vowel) *t*-type words. Fundamental frequency, onset, and subjective loudness of the signals were equalized. The duration, could be varied in 8 msec steps. The used vowels were

[oe], [ɔ], [ɑ], [u], [y], [i], [a], [ɸ], [o], [ɛ], and [e]

(IPA, 1967). A more extended description of the experimental setting is given in POLS et al. (1969). Fifteen male subjects, all with normal hearing, participated in the identification experiment. Before the actual experiment started, the subjects first had the opportunity to become familiar with the stimuli. To this purpose a complete series of 11 stimuli was presented to the subject who was instructed to write down on an answer sheet next to the numbers from 1 to 11 how he would name the sound. In fact, all subjects used vowels for names. The series was repeated until in the opinion of the experimenter a subject had a sufficiently high percentage of correct identifications. Thereafter the main experiment started in which the subject was instructed to respond with the number of the stimulus

corresponding to the name given during the last familiarization series. The two experimental conditions in the main experiment were one with stimulus duration of 1 period (8 msec) and another with 32 periods (256 msec). Under both conditions the subject was presented with 6 sequences of the 11 stimuli all 66 placed in random order. The percentages correct answers for the 32 period situation varied per subject from 54.5 % to 94.0 % with a mean value of 70.4 %.

Misidentifications between related short-large vowels like the pairs [œ]-[ø], [ɔ]-[o] and [ɑ]-[a] accounted for 14.0 % of the errors. As was expected, the 32 period confusion matrix showed many empty cells. For this reason, most of the techniques that can be used for the analysis of the 1 period confusion matrix were not applicable. Therefore, the analysis of the 32 period matrix will not be reported in this study.

### 3. ANALYSIS OF THE CONFUSIONS

The  $11 \times 11$  confusion matrix for the stimulus duration of 1 period is given in table 1. Entry ( $i, j$ ) in this matrix gives the number of times stimulus  $i$  is identified as  $j$ . The percentage correct answers varied from 7.5 % to 51 %, with a mean of 30.8 %, given a prior probability of 9 %.

TABLE 1

Matrix of errors in identifications (confusion matrix). Entry ( $i, j$ ) gives the number of times stimulus  $i$  is identified as  $j$ .

	[œ]	[ɔ]	[ɑ]	[u]	[y]	[i]	[a]	[ø]	[o]	[ɛ]	[e]
	u	o	a	oe	uu	ie	aa	eu	oo	e	ee
	1	2	3	4	5	6	7	8	9	10	11
1	28	2	4	7	6	1	0	17	11	6	8
2	8	31	10	10	0	3	0	4	4	10	0
3	1	6	23	5	3	6	21	7	7	7	4
4	4	8	5	50	4	0	1	2	8	5	3
5	11	4	2	12	26	11	1	7	1	2	13
6	3	0	2	3	14	56	2	4	0	0	6
7	2	1	22	0	4	12	35	1	2	4	7
8	26	1	1	2	10	6	1	22	2	9	10
9	5	18	18	5	3	5	11	6	10	6	3
10	2	8	14	2	3	10	18	2	8	15	8
11	22	6	5	3	10	6	3	13	4	9	9
Totals	112	85	106	99	88	116	93	85	67	73	71

How can the confusion errors, given in table 1, be analyzed in such a way that a configuration of metric distances and the coordinates of the stimuli are obtained? One class of methods is applied directly to the asymmetric confusion matrix; therefore we will call them direct methods of analysis. The other class of methods consists of three different techniques, for two of them a symmetry is achieved by a simple averaging process and a third reconstructs a symmetric matrix by a special correction for response bias. Thereafter, Kruskal's multidimensional scaling procedure for symmetric matrices was used in the analysis of our data (KRUSKAL, 1964a,b).

### 3.1. *Direct methods of analysis*

An inspection of the raw confusion matrix (table 1), immediately reveals its asymmetry. The largest differences that occur are between the pairs [oe]–[e], [ɔ]–[o] and [a]–[e]. Taking these rather large differences into consideration, it becomes questionable whether symmetrizing the matrix by some averaging procedure is justified. Therefore we shall first look at the results of an analysis which does not require a symmetric matrix.

The first technique used was developed by DE LEEUW (1968). It is a nonmetric multidimensional scaling program in the series for the canonical analysis of relational data (CDARD-7). The data obtained in our scaling experiment can be described in terms of a nonmetric multidimensional scaling model as a set of quadruples of indices. Let  $S$  be the set of stimuli. The fact that a given quadruple  $(i, j, k, l)$  is in the data set  $L$  means that the dissimilarity of  $s_i$  and  $s_j$  is greater than the dissimilarity of  $s_k$  and  $s_l$ . The objective of this nonmetric multidimensional analysis is to map the stimuli in set  $S$  into a Euclidean  $p$ -dimensional space with metric  $d$  in such a way that  $(i, j, k, l) \in L \Rightarrow d_{ij} - d_{kl} \geq 0$ , where  $d_{ij}$  and  $d_{kl}$  are the distances between stimuli  $i$  and  $j$ , and between stimuli  $k$  and  $l$  respectively. It is desired that the data will be represented parsimoniously in the  $p$ -dimensional coordinate space. The dimensionality  $p$  is sought to be as small as possible, satisfying the monotonicity condition that  $d_{ij} > d_{kl}$  whenever the observed data indicate that  $s_k$  is 'closer' to  $s_l$  than  $s_i$  to  $s_j$ . It is assumed that no more than one pair of stimuli is identical and that in all cases  $i \neq j$  and  $k \neq l$ . Several algorithms can be used in order to reach this goal (see DE LEEUW, 1968). As initial configuration a solution is obtained by computing principal components. This CDARD-7 program yields two solutions, a row (or stimulus) solution, and a column

(or response) solution. How good the obtained CDARD-7 solution is cannot be evaluated by a measure of goodness of fit that is comparable with Kruskal's stress.

In the CDARD-7 procedure the initial number of elements, i.e. the number of initial binary relations, in the data set  $L$  is computed. Given  $n$  stimuli, the number of binary relations would be  $(1/8)n(n-1)(n-2)(n+1)$ . In our case, however, because of equal cells in the confusion matrix (most of them are empty) the number of initial binary solutions is 451. As an indication of the appropriateness of a certain solution the number of violations has been taken. For a 3-dimensional solution the number of violations of the initial 451 binary relations becomes 74. This seems quite satisfactory.

The next step is to match the resulting configuration with the target. A procedure for matching one configuration with another given configuration can be described as follows. For a matrix  $A$  an orthogonal transformation matrix  $T$  is sought such that  $AT = \hat{A}$  is as similar as possible to the given target matrix  $B$ . The criterion for fit is maximizing the covariance between corresponding points on corresponding axes; this is equal to minimizing the sum of squares of distances between corresponding points. This can be obtained by maximizing the function.

$$\varphi = \sum_{i=1}^n \sum_{k=1}^r a_{ik} b_{ik} \quad \begin{array}{l} i=1, \dots, n \text{ stimuli} \\ k=1, \dots, r \text{ dimensions} \end{array}$$

The fact that  $A$  and  $B$  must be as similar as possible can also be formulated in terms of the difference of  $A$  and  $B$ , being the matrix  $E$ . So,  $AT - B = E$ . Minimizing  $\text{tr}(E'E)$  leads to an  $A$  that is as similar as possible to matrix  $B$  in a least squares sense. For a detailed discussion of the possible analytical techniques for solving the problem of fitting a target we refer to CLIFF (1966), SCHÖNEMANN (1966) and HORST (1965). As a measure of the correspondence between the rotated configurations and the target configuration, product moment correlation coefficients between the projection of the stimuli on the corresponding dimension of both configurations can be used. If both configurations are normalized these correlation coefficients are identical with Tucker's coefficient of congruence. The correlation coefficients along the corresponding dimensions are 0.851, 0.924 and 0.862 for the response solution and 0.892, 0.772 and 0.597 for the stimulus solution (table 2). Except perhaps for the third dimension of the stimulus configuration the matching of the obtained

TABLE 2

Correlation coefficients for the configurations, recovered by the different techniques, rotated to maximal congruence with the perceptual configuration from similarity judgements.

Techniques used	Dimensions		
	I	II	III
CDARD-7 column (response) solution	0.851	0.924	0.862
CDARD-7 row (stimulus) solution	0.892	0.772	0.597
UNFOLD column solution	0.941	0.804	0.728
UNFOLD row solution	0.784	0.919	0.816
KRUSKAL solution with correction for response bias	0.980	0.989	0.991
KRUSKAL solution using a measure of stimulus generalization	0.945	0.903	0.761
KRUSKAL solution using simple averaging procedure	0.933	0.925	0.837

CDARD-7 configurations with the perceptual configuration from similarity judgements is rather good.

Another method for analyzing conditional proximity matrices is a program called UNFOLD designed by ROSKAM (1968). This is in fact a computerized iterative procedure (using Kruskal's algorithm) of Coomb's unfolding theory based on a distance model. The problem in general is to represent  $n$  objects geometrically by  $n$  points, so that the interpoint distances correspond in some sense to experimental similarities between objects; the fundamental hypothesis for these methods of analysis being that dissimilarities are monotonically related with interpoint distances. Using Kruskal's algorithm (KRUSKAL, 1964b) a solution for this problem can be obtained with a stress value as a measure of goodness of fit. In the UNFOLD procedure for each row of the raw confusion matrix the rank order of similarities is determined; this is also done for each column. So, monotonicity is constrained within rows or columns only, leading to a row (stimulus) and a column (response) solution.

Instead of looking at the rank order of similarities within rows and columns separately another procedure called CONPROX can be used for analyzing conditional proximity data. In this approach the presented stimuli and the identifying responses are considered as the elements of one set. For every element of this set the rank order of (dis)similarity is determined in relation to the other elements of the set. Multidimensional

scaling approach for nonsymmetric matrices based on such a procedure might give rise to distortions (depending upon the degree of asymmetry) and eventually to psychologically unacceptable treatments (e.g., if such a procedure is used for the analysis of sociometric data). These difficulties might appear for the simple fact that no distinction is made between stimuli and the identifying responses, thus resulting in a single configuration. The psychological interpretation of a single configuration, however, is much easier than that for two solutions, one for the columns and another for the rows. If there is a substantial difference between the row and the column solutions and if this discrepancy is real, then this must have implications for the theory of perception of speech sounds. Our experiment, however, is too specific to get more insight into this problem. Therefore, we will leave this problem for what it is.

The CONPROX program resulted in a single 3-dimensional configuration with a stress of 25.3 %, which is, according to the indications given by ROSKAM (1968) poor. Matching this configuration with the target configuration (i.e., the perceptual space from similarity judgements) also gave unsatisfying results. So this solution was rejected. The UNFOLD procedure yielding a stimulus and a response solution, was quite good. A 3-dimensional configuration resulted in a minimum stress of 0.94 %. Rotation of the row and column solutions to maximal congruence with the target configuration resulted in configurations for which the coefficients of correlation between the corresponding dimensions of the rotated structure and the target structure are 0.942, 0.804, 0.728 and 0.784, 0.919, 0.816 for the 3 dimensions of the response and stimulus respectively (table 2). The correspondence between these solutions from identification data and the solution from similarity judgements is as good as in the case with the CDARD-7 solutions. It remains to be seen, however, whether it is necessary to obtain separate solutions for the stimuli and for the identifying responses. Furthermore, if response bias occurs in the process of identification of the stimuli, i.e., if some responses are given more frequently than other ones, then the described procedures are inadequate.

### 3.2. *Indirect methods of analysis*

From table 1 it is clear that the observed confusion matrix is not symmetric. This might be due to the fact that the subjects tend to use some responses more than others. In order to correct for response bias a method was used described by WAGENAAR (1968). All possible  $2 \times 2$

matrices are constructed from the raw confusion matrix in such a way that the proportions remain the same. Each  $2 \times 2$  biased confusion matrix then can be considered as the product of a symmetric matrix and a response bias matrix. The decomposition procedure is as follows. Let such  $2 \times 2$  biased confusion matrix be of the following form

$$\begin{array}{rcc} & & \text{response (receiver)} \\ & & \begin{array}{cc} i & j \end{array} \\ \text{stimulus} & \begin{array}{c} i \\ j \end{array} & \begin{bmatrix} A - \lambda_{ij}A & B + \lambda_{ij}A \\ B - \lambda_{ij}B & A + \lambda_{ij}B \end{bmatrix}, \\ \text{(giver)} & & \end{array}$$

where  $\lambda_{ij}$  stands for the response bias factor. This biased confusion matrix, in general nonsymmetric, can be decomposed as follows:

$$\begin{bmatrix} A - A\lambda & B + A\lambda \\ B - B\lambda & A + B\lambda \end{bmatrix} = \begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{bmatrix} 1 - \lambda & \lambda \\ 0 & 1 \end{bmatrix},$$

in which the first matrix on the right hand side is a symmetrical one and the second a response bias matrix. By solving for the response bias factor the  $2 \times 2$  symmetrical unbiased matrix can be constructed. This can be done for all possible pairs of stimuli from the  $11 \times 11$  raw confusion matrix. However, such a procedure presupposes that it is justified to split the  $11 \times 11$  matrix in all possible  $2 \times 2$  matrices. Or, to say it in terms of discriminations among stimuli, that a subject's discrimination between stimuli  $i$  and  $j$  is independent of the set of stimuli from which  $i$  and  $j$  are taken. The probabilistic analogue of this statement of 'independence from irrelevant alternatives' is the requirement that the ratio of the probability of confusing two stimuli should not depend on the total set in which the two stimuli are embedded. This is equivalent to Clarke's constant ratio rule (CRR) stating 'that the ratio between any two entries in a row of a submatrix is equal to the ratio between the corresponding two entries in the master matrix' (CLARKE, 1957). Actually, the CRR is the same as Luce's choice axiom (LUCE, 1959). There is ample evidence that the CRR holds for speech sounds. In several experiments the CRR was evaluated (CLARKE, 1957, 1959; HODGE and POLLACK, 1962; POLLACK and DECKER, 1960). In general, it can be concluded that the ratio of confusion errors among any two items of the message-set is undisturbed by removal or addition of other items. After correcting for response bias the  $11 \times 11$  symmetric confusion matrix and the  $11 \times 11$  bias matrix are obtained (table 3).

TABLE 3

Entries above the main diagonal are the elements of the matrix of response bias. Entries below the main diagonal are the elements of the dissimilarity matrix corrected for response bias (scaled from 0 to 100).

	1	2	3	4	5	6	7	8	9	10	11
1		0.14	-0.08	-0.13	-0.12	0.02	0.05	0.16	0.05	-0.06	0.49
2	92.27		-0.04	-0.11	0.13	-0.09	0.03	-0.07	0.33	0.10	0.40
3	95.45	78.52		-0.09	-0.04	-0.17	-0.09	-0.19	0.41	0.25	0.21
4	91.53	84.57	90.04		0.24	0.05	-0.02	0.04	0.10	0.03	0.19
5	79.93	100.0	92.53	90.23		-0.10	0.07	0.10	0.19	0.10	0.19
6	96.69	100.0	95.83	100.0	77.84		0.22	0.15	0.33	0.40	0.30
7	100.0	100.0	57.53	100.0	96.04	95.57		0.02	0.47	0.44	0.08
8	54.82	95.32	94.63	95.97	76.42	92.18	97.18		0.29	-0.27	0.28
9	70.27	53.44	60.48	84.67	95.41	100.0	89.81	88.24		-0.03	0.08
10	87.50	72.78	68.91	90.06	92.11	100.0	81.59	85.78	64.25		0.15
11	56.64	100.0	81.27	92.98	58.70	86.11	81.82	56.69	75.00	58.97	

Then the multidimensional analysis of the symmetric matrix was performed by the Kruskal-program. A solution in three dimensions was chosen on the basis of prior results from experiments on the perceptual structure of these vowels and by considering the stress-values for various solutions. The obtained 3-dimensional configuration with a stress of 2.9 % was then rotated to maximal congruence with the target matrix. The congruence between the rotated configurations and the target is excellent as is indicated by the correlations between the projections of the stimuli of the corresponding configurations for the 3-dimensions, being 0.908, 0.989, and 0.991 respectively (see table 2).

Another method for removing the asymmetries is one based on a model for stimulus and response generalization by SHEPARD (1957, 1958). Shepard suggested the following formula

$$s_{ij} = \left[ \frac{(p_{i/j}) p_{(j/i)}}{p_{(i/i)} p_{(j/j)}} \right]^{\frac{1}{2}},$$

where  $s_{ij}$  indicates the similarity between stimuli  $i$  and  $j$ , and  $p_{(i/j)} \neq p_{(j/i)}$  the probability with which stimulus  $i$  (or  $j$ ) is identified as  $j$  (or  $i$ ). Shepard's measure for stimulus generalization, which equals Luce's measure of stimulus similarity (LUCE, 1963), can be regarded as an averaging technique for removing the asymmetries in the original confusion matrix. In fact the geometric mean is taken from the two corresponding

probabilities of misidentifications, taking the corresponding correct identifications into account. Instead of using the geometric mean we can also calculate a simple arithmetic mean. This leads to the following formula:

$$s_{ij} = \left[ \frac{p(i/j) + p(j/i)}{p(i/i) + p(j/j)} \right],$$

where the elements are defined as above for Shepard's measure.

Both similarity measures based on averaging were used to symmetrize the raw confusion matrix. Both results were analyzed by the Kruskal program and thereupon rotated to maximal congruence with the perceptual structure obtained earlier (table 2).

There is a striking resemblance of the two last mentioned similarity measures and the two distance measures WILSON (1963) used in multidimensional analysis of English consonants. Because Wilson analyzed his data by Torgerson's multidimensional scaling technique distance measures had to be defined for transforming frequencies of confusion errors into distances. We, however, used Kruskal's technique which makes it unnecessary to define a distance measure in advance. This scaling procedure makes less assumptions than Torgerson's method.

Comparing the three rotated configurations, resulting from the three described symmetrizing procedures prior to multidimensional scaling with the target structure, it can be seen that the solution obtained by applying Luce's choice axiom and Wagenaar's correction for response bias yields the highest correlation coefficients. In general, there is a good agreement among the final configurations as can be seen from fig. 1 (dimension I versus dimension II), and fig. 2 (dimension I versus dimension III) in which the stimuli are plotted with coordinates according to the rotated configurations and the target structure.

#### 4. DISCUSSION AND CONCLUSION

The main purpose of this study was to investigate whether it was possible to recover a perceptual configuration from an experiment on vowel identification which is in some way related to a psychological space constructed from similarity judgements. No attempt is made to relate the results to physical variables nor to the phonemic space of these vowels (see POLS et al., 1969). Of the direct methods of analysis the CDARD-7 and the UNFOLD-program gave interesting results. The resulting row as well as column solutions are very similar to the per-

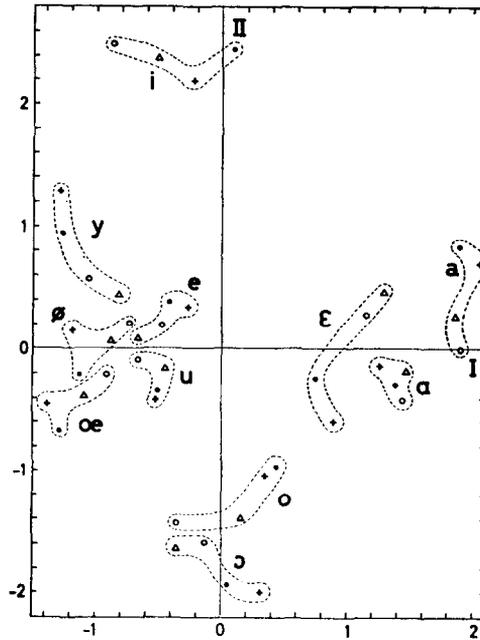


Fig. 1. Dimension I versus dimension II for the three matched configurations with the target structure.

- △ points according to the method correcting for response bias,
- + points according to the arithmetic mean procedure,
- points according to the measure for stimulus generalization,
- points according to the target structure.

ceptual similarity structure, suggesting that the two solutions are not essentially different. Are the solutions given by the just mentioned programs the most adequate solutions? First of all there is no statistical measure for the adequacy for such solutions. Secondly, it is not quite clear how to interpret the column and the row solutions in view of an available theory of speech perception. Both do not differ systematically from each other, nor from the perceptual structure obtained by similarity judgements. Besides that it is implicitly assumed in both procedures that no response bias occurred. If it does, the procedures based on a distance model are inadequate (ROSKAM, 1968). Looking at the identifying responses given by the subjects a tendency can be noted that some res-

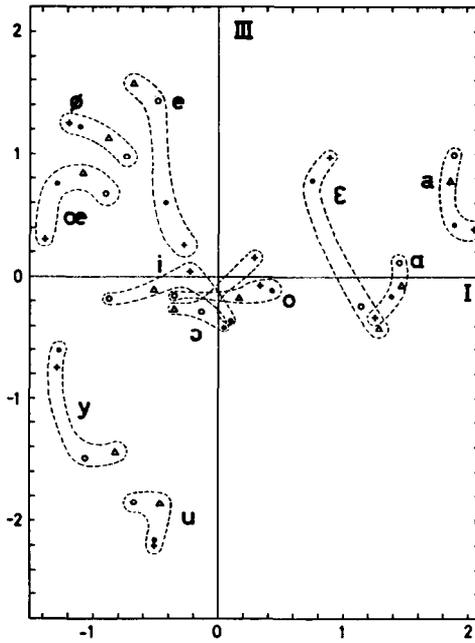


Fig. 2. Dimension I versus dimension III for the three matched configurations with the target structure. For explanation of the points see fig. 1.

ponses are used more than others, suggesting that response bias does occur.

Of the indirect methods of analysis the solution based on Luce's choice axiom to correct for response bias is best in terms of agreement with the target structure, being the similarity structure. Shepard's measure for stimulus generalization and the arithmetic averaging procedure also yield solutions that are in close agreement with it, but the fit for the third dimension is not as good as with the first mentioned method. All these techniques, however, recover configurations which have the same structure and which can be interpreted in the same way as the perceptual space of similarity judgements of the vowels. In spite of the fact that symmetrizing by averaging might not be justified (see e.g. COOMBS, 1964) it does lead to solutions which seem in better agreement with the target. We must keep in mind, however, that the 'averaging' techniques result in similarity matrices in which the sum of the different rows (or columns) are not equal. It seems reasonable to conclude that

an appropriate theoretical analysis (i.e., Luce's choice axiom and the correction for response bias) that extracts a symmetric similarity matrix must be preferred to the averaging procedures.

(Accepted September 15, 1970.)

#### REFERENCES

- CLARKE, F. R., 1957. The constant-ratio rule for confusion matrices in speech communication. *Journal of the Acoustical Society of America* **29**, 715-720.
- , 1959. Proportion of correct responses as a function of the number of stimulus-response alternatives. *Journal of the Acoustical Society of America* **31**, 835 (abstract).
- CLIFF, N., 1966. Orthogonal rotation to congruence. *Psychometrika* **31**, 33-42.
- COOMBS, C. H., 1964. *A theory of data*. New York: Wiley.
- DE LEEUW, J., 1968. *Non-metric multidimensional scaling. RN 010-16*. Psychological Institute University of Leyden, Leyden.
- HODGE, M. H. and I. POLLACK, 1962. Confusion matrix analysis of single and multidimensional auditory displays. *Journal of Experimental Psychology* **63**, 129-142.
- HORST, P., 1965. *Factor analysis of data matrices*. New York: Holt, Rinehart and Winston.
- IPA, 1967. *The principle of the international phonetic association* (reprinted). IPA, Department of Phonetics, University College, London, W.C. 1.
- KRUSKAL, J. B., 1964a. Multidimensional scaling by optimizing goodness of fit to the nonmetric hypothesis. *Psychometrika* **29**, 1-27.
- , 1964b. Nonmetric multidimensional scaling: A numerical method. *Psychometrika* **29**, 115-130.
- LUCE, R. D., 1959. *Individual choice behavior*. New York: Wiley.
- , 1963. Detection and recognition. In: R. D. Luce, R. R. Bush and E. Galanter, (eds.), *Handbook of mathematical psychology*. New York: Wiley, Vol. I, chap. 3, pp. 103-189.
- MILLER, G. A., 1956. The perception of speech. In: M. Halle et al. (eds.), *For Roman Jakobson: essays on the occasion of his sixtieth birthday*. The Hague: Mouton, pp. 353-359.
- POLLACK, I. and L. DECKER, 1960. Consonant confusions and the constant-ratio rule. *Language and Speech* **3**, 1-6.
- POLS, L. C. W., L. J. TH. VAN DER KAMP and R. PLOMP, 1969. Perceptual and physical space of vowel sounds. *Journal of the Acoustical Society of America* **45**, 458-467.
- ROSKAM, E. E. CH. I., 1968. *Metric analysis of ordinal data in psychology*. Thesis, University of Leyden.
- SCHÖNEMANN, P. A., 1966. A generalized solution of the orthogonal Procrustes problem. *Psychometrika* **31**, 1-10.

- SHEPARD, R. N., 1957. Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. *Psychometrika* **22**, 325-345.
- , 1958. Stimulus and response generalization: deduction of the generalization gradient from a trace model. *Psychological Review* **65**, 242-256.
- TORGERSON, W. S., 1958. *Theory and methods of scaling*. New York: Wiley.
- WAGENAAR, W. A., 1968. Application of Luce's choice axiom to form discrimination. *Nederlands Tijdschrift voor Psychologie* **23**, 96-108.
- WILSON, K. V., 1963. Multidimensional analysis of confusions of English consonants. *American Journal of Psychology* **76**, 89-95.