ANALYSIS OF ACOUSTIC INVARIANCE IN THE PLACE OF ARTICULATION OF STOPS BY MEANS OF THE MDS

Josep Matas-Crespo*, Eugenio Martínez-Celdrán & Alicia Ortega-Escandell
Universitat de Barcelona. Spain.

ABSTRACT

Locus equations allow us to maintain that the regression slopes in the correlation of the start of the F2 and the middle of the same F2, which discriminate the places of articulation of stops perfectly, although it does not take into account the slope calculated for each token. MDS ("MultiDimensional Scaling"), provides a new perspective of acoustic invariance among the stops. This statistical technique offers a graphic representation of the objects in a geometric space. In this experiment, the values at the start and the middle of the F2 have been used to calculate the slope of each token. With these values a matrix of proximities among the different stops was obtained. By using this matrix, the MDS has been able to construct a space that gathers in a same zone the couples voiceless-voiced and separates in different fields the three places of articulation. The results have been discussed in terms of the progress that this technique represents compared to the locus equations approach.

1. INTRODUCTION

The discrimination levels of the P.A. of stops based on the slope of each token has never reached more than 77%. This relative failure led some researchers to the development of a recent approach to the problem: locus equations. This methodology has provided highly successful discrimination of stops (100%), taking as discriminant variables the slope and the y-intercept in the locus equations [1]. This paper aims to demonstrate that the value of the slope calculated on each of the tokens allows us to discriminate among the stops as satisfactorily as the parameters in the locus equations. The estimation of discriminant functions among *k* experimental groups will be hindered if such parameters do not fulfill the requirements of minimal variability, distributional normality and homocedasticity. Contrary to that, a statistical technique that does not rely on such distributional restrictions will be optimal to discriminate the groups. This is the case of Multidimensional Scaling (MDS), a statistical technique the origin of which can be found in the works of Torgerson [3], which was substantially developed in the paper by Kruskall [4]. The main goal of the MDS is the construction of a metric with the minimum possible number of dimensions that allow us to represent the proximities among the elements with the highest degree of accuracy ([5]). An application of this technique to the study of the Catalan vowels can be found in J. Matas ([6]).

2. METHOD

2.1. Informants: 10 informants (5 male and 5 female) aged between 18 and 30.

2.2 Procedure: Each of the informants produced 150 sequences with the structure [kan'CVna], where C corresponded to one of the Castilian Spanish stops ([p,t,k,b,d,g]) and V to one of its vowels ([i,e,a,o,u]). Each of the possible combinations was repeated 5 times. The utterances were recorded in a soundproof booth by means of a "Shure SM58" microphone and a Marantz (model CP430) tape recorder.

2.3 Acoustic analysis: The informants' utterances were analyzed using the CSL 4300 B program (by Kay Elemetrics). Two values of the vocalic F2 were obtained by cursor positioning on a wideband spectrogram on which the values of the LPC derivation of the formants was superimposed (using the "Formant History" option). The first value (F2i) corresponded to the first visible glottal pulse after the burst. The second value (F2v) was taken at the middle of the vocalic static period. The slope was calculated using the formula (F2i-F2v)/ΔT (where T stands for "time" and Δ for the "increment").

3. RESULTS

Table 1 shows the average values and standard deviations of the value of the slope for each place of articulation and vowel, as well as the contrast of homogeneity of variances by means of Levene's test. As we can see, the deviation values are too large, and there is no guarantee that in such cases the variances are equal. With such distributional conditions it is difficult that a Discriminant Analysis (D. A.) provides discriminant functions with a high predictive power. However, a technique such as the MDS should be more robust in such conditions, since its success does not depend on normal distributions. To verify this hypothesis two analysis were carried out: a D.A. within each vowel, taking as discriminant variables the value of the slope and the y-intercept for each token, and an MDS only with the slope values and not including the influence of the vowel. Not surprisingly, the

* Author to whom all correspondence should be addressed: Laboratori de Fonètica, Gran Via de Les Corts Catalanes, 585 08007 Barcelona –Spain.
E-mail: matas@lingua.fil.ub.es

D.A. did not reach an average percentage of discrimination for the five vowels of more than the 71%.

Table 1

	Pend	labial	dental	velar	Levene	prob
a	Mean	,30639	-,04854	-,23556	22,4	0
	SD	,12420	,12493	,23033		
e	Mean	,47402	,35110	-,05380	7,42	0,001
	SD	,14023	,19683	,24207		
i	Mean	,46546	,44214	,04215	1,53	0,216
	SD	,21392	,22070	,20010		
o	Mean	,08222	-,36545	,01215	1,98	0,139
	SD	,12525	,12563	,14996		
u	Mean	,00267	-,60287	,02169	3,61	0,028
	SD	,20608	,26332	,24002		

3.1 Analysis by means of the MDS

This technique is based on the assumption that, for a group of elements, we can obtain a matrix of experimental proximities from which to create a t-dimensional space (where t is equal to the number of dimensions) and the elements are represented in such a way as to ensure that the distances among the obtained points are equivalent to the proximities experimentally found. The minimum condition of equality establishes that the rank order of the proximities must be equal to the rank order of the distances among the points in the configuration. It must be pointed out, too, that the work by Young, Takane & Lewicky [7], where the algorithm ALSCAL based on the method of the *alternating least squares* was developed, improves computationally the creation resultant configuration. In addition, this technique offers the advantage of making it possible to represent graphically the objects in a geometric space. In order to have an easily understandable graphic representation, the number of dimensions constituting this space should not be more than three.

3.2 Matrix of similarities: The matrix of similarities was obtained from the formula of the Euclidean distance, which is defined as the square root of the summatory of the differences of the elements to the square power:

Table 2

	B	D	G	K	P
B	,000				
D	6,764	,000			
G	7,225	8,426	,000		
K	6,754	8,012	4,066	,000	
P	2,776	7,147	7,954	7,552	,000
T	6,779	3,599	7,712	7,266	7,219

The ALSCAL algorithm was applied to this matrix (Table 2) of distances. In this case the computation was carried out with the help of the statistical package SPSS 8.0. The restrictions imposed on the computations were that the convergence criterium (Improvement) was under 0,0001 or else that the Stress reached a minimum value of 0,0005. Four iterations sufficed for this first analysis. Figure 1 shows the configuration obtained with the values of adjustment (Stress)[1], as well as the percentage of explained variance (RSQ) for this configuration.

Figure 1 (end of document) shows that the distance between the pairs /b/-/p/; /d/-/t/; and /k/-/g/ is minimal compared to the distance between the three pairs. Dentals appear at the bottom left corner; velars at the bottom right and labials at the top left corner. A straightforward conclusion from such a configuration is that place of articulation can be successfully represented from the computed distances, using only, for each token, the slopes of the second formant of the postconsonantal vowel in CV structures.

4. CONCLUSIONS

We can affirm that the values of the slope obtained for each token contain a series of distributional properties that make them unsuitable for a statistical treatment by means of a technique that requires certain conditions for its application. The MDS has turned out to be a highly satisfactory technique for the treatment of the slopes of each token, providing as a result a configuration that has proved capable of neatly differentiating the three places of articulation of Spanish stops.

5. DISCUSSION

Two basic questions arise from these results: To what extent do locus equations involve an advance? Are they the only plausible approach to the transitions of the F2 associated to the stop place of articulation? In the first place, we must consider what a locus equation really is. The locus equation is an adjustment based on the least squares for *n* coordinates, in which the *y* axis is the F2 in at the vocalic start and the *x* axis is the F2 at the static part of the vowel, where all the vowels and each of the subjects are separately considered. After all, it involves a measure of the average increment and variance of F2 as expressed by two coefficients: the slope of the equation *b* and the y-intercept *a*. Locus equations successfully manage to dismiss variations due to chance. Furthermore, when the equation is obtained for each informant separately, a normalization effect on the variability the informant could generate is produced. [9] The final product is a series of new measures (*a*,*b*), for each informant, which do discriminate the stops, but tell us very little about the true increment of the F2. It is in this sense that the MDS involves an advance with respect to locus equations, since the kind of metric it works with is the real increments (which takes into account the temporal period over which the increment takes place) expressed as a value of the slope. These increments are used to express the proximities, for example, between a [b] and a [d], or a [b] and a [p], (see Table 2). It is from the combinations of all these proximities that we can obtain a

[1] The optimal Stress values can be found in Spence and Ogilve [8]

space where the three places of articulation remain clearly differentiated. These results allow us to vindicate again the importance of the slope calculated in each case as an acoustic cue associated to the place of articulation of stops. This opens up a new perspective. The question that needs to be tackled with now is to discover up to what point the MDS can be used as a computational algorithm for automatic speech recognition. Nonetheless, the results of this experiment need to be replicated in other studies, which will confirm its descriptive and classificatory power.

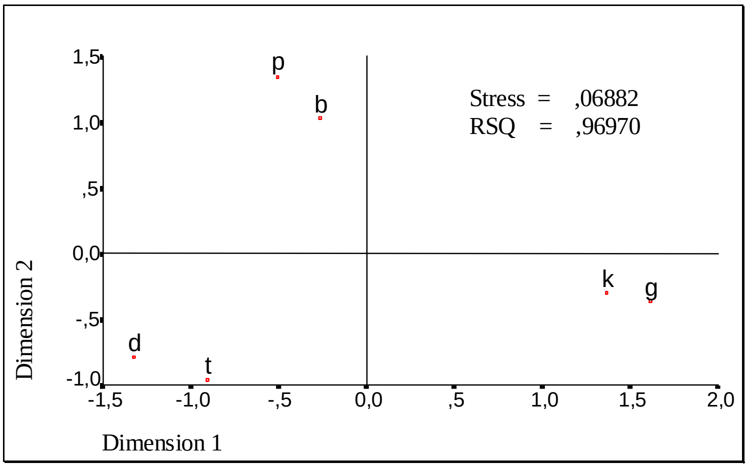

Figure Error! Unknown switch argument.

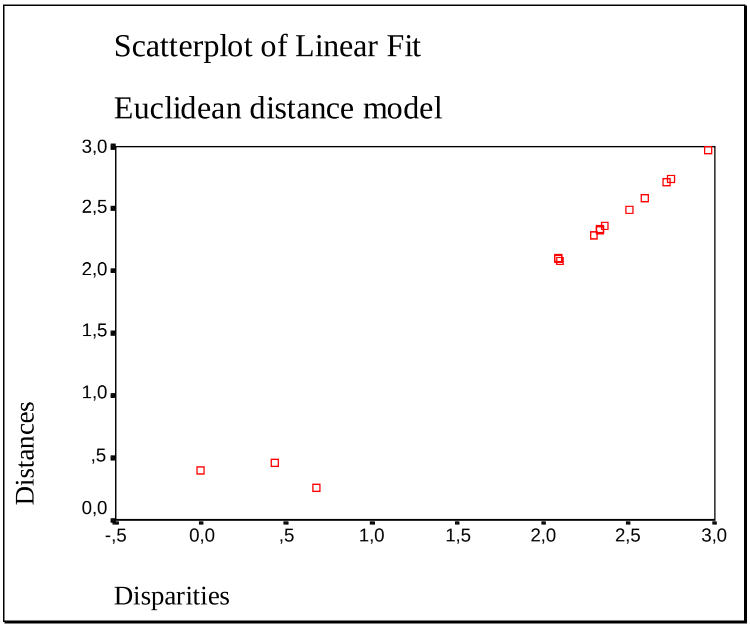

Figure Error! Unknown switch argument.

REFERENCES

[1] Sussman,, H. M., MacCaffrey, H. A., and Matthews, S. A. 1991. An investigation of locus equations as a source of relational invariance for stop place categorization. *J. Acoust. Soc. Am.* 90, 1309-1325

[2] Celdrán (Martínez-Celdrán), E. and Villalba, S. (1995) Las ecuaciones de locus y el punto de articulación del español. *Estudios de Fonética Experimental* vol. VII, 86-109

[3] Torgerson, (1952) *Multidimensional Scalling: I. Theory and Method*. Psychometrika,17, 401-419.

[4] Kruskal, J.B. (1964) Multidimensional Scalling by optimizing goodness of fit to a nonmetric hypotesis. *Psychometika*, 29, 1-27.

[5] Davison, M.L. (1983) *Multidimensional scalling.* Nueva York. John Willey.

[6] Matas-Crespo, J.Validació de la carta vocàlica del català oriental central mitjançant la tècnica de l'escalament multidimensional. *Estudios de Fonética Experimental* vol. VIII, 254-270
[7] Takane, Y., Forrest, W., Leeuw, J. 1977. Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika.* vol 42, 1 7-67
[8] Spence, I., and Ogilve, J.C. (1973) A table of spected stress values for random in multidimensional scaling. *Multivariate Behavioral Research*, 8, 511-518.
[9] Sussman M.H., Bessell, N., Dalston, E., and Majors, T. 1997. An investigation of stop place of articulation as a function of syllable position: A locus equation perspective. *J. Acoust. Soc. Am.* 101 (5) 2826-2838