THE VOWEL SYSTEM OF ITALIAN CONNECTED SPEECH

Federico Albano Leoni, Francesco Cutugno, Renata Savy
CIRASS - Università di Napoli "Federico II"

ABSTRACT

A research on the spectro-acoustic features of Italian vowels in connected speech will be presented. The material used for our analyses was recorded from four regional TV news bulletins. 6400 vowels were analysed (40 speakers uttering 10 tokens of 16 vowel types).

We will present some results relative to the male half of the corpus. We will focus on some aspects such as reduction and centralization, overlaps between adjacent vowels, vowels duration.

All our data are available via Internet, see note in the last page after References for details.

INTRODUCTION

The earliest spectroacoustic description of the Italian vowel system was made by Ferrero in 1968 [1].

The data presented at that time were obtained from a corpus of vowels pronounced in isolation by 20 speakers native of northern regions of Italy. All the following attempts to describe the Italian vowel system suffered from the same limitations [2], [3], [4].

For the present work we decided, on the contrary, to create a phonetic database based on recordings of TV news bulletins (see also [5], [6]).

In our view this kind of speech material represents an example of standard Italian of middle-high level and, at the same time, an example of connected speech which, although partly read by the journalists, can be seen as "spontaneous".

METHODS

The total corpus consists of 6400 items taken from recordings of regional bulletins broadcast by the national TV company (RAI) in four regions of Italy: Lombardy, Tuscany, Latium and Campania. For each region 10 speakers were chosen (5 males and 5 females); from each of the total 40 speakers 10 tokens were used of 16 different vowel types (stressed [i, e, ɛ, a, ɔ, o, u], unstressed non final [i, e, a, o, u], unstressed final [i, e, a, o]). The tokens were extracted from "full words" (nouns, adjectives, verbs, adverbs) excluding diphthongs. We organized our data into a database containing the following information:

-speaker id.;
-word uttered;
-stress conditions and position of the vowel (stressed, unstressed, final);
-preceding and following segments;
-preceding and following vowels;
-syllable structure (open or closed);
-total vowel duration;
-maximum pitch within the vowel;
-maximum energy within the vowel;
-f1, f2, f3 values measured from average FFT spectrum of the central portion (second third) of the vowel.

Presently (April 95), only data from the male speakers portion of the corpus are completely available, while the analysis of the female portion is going to be completed within a few months.

RESULTS

General results

The average and σ values for f1 and f2 are presented in Table 1.

	i		e		ɛ		a		ɔ		o		u		
	f1	f2	f1	f2	f1	f2	f1	f2	f1	f2	f1	f2	f1	f2	
stressed	273	2234	375	2037	493	1855	702	1488	536	1043	404	983	307	895	mean
	54	173	62	205	76	163	91	112	66	147	62	164	57	145	σ
unstress.	283	2099	410	1775			564	1454			433	1091	315	1013	mean
	54	198	77	217			94	176			79	167	65	213	σ
final	294	2106	435	1790			567	1497			438	1134			mean
	54	212	94	176			85	142			77	172			σ

*Table 1. Mean values and σ of f1 and f2 (Hertz) for **stressed, unstressed** non final, and unstressed **final** vowels.*

In Figures 1, 2 and 3 the data of Table 1 are presented in terms of f1/f2 diagrams on a Hertz linear scale. In these figures each ellipse defines the distribution area of a vowel. The coordinates of the ellipse centers are the mean values of f1 and f2 while the axes are±1σ.

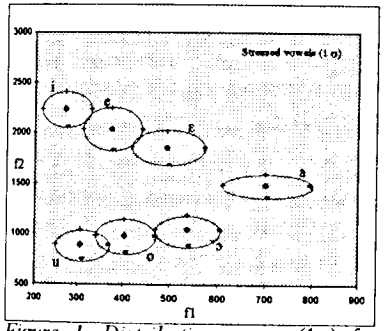

Figure 1. Distribution areas (1σ) for Italian stressed vowels.

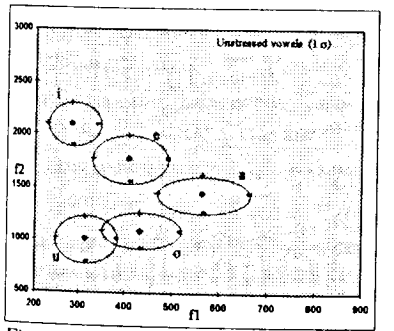

Figure 2. Distribution areas (1σ) for Italian unstressed non final vowels.

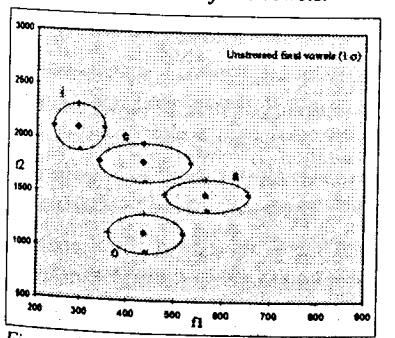

Figure 3. Distribution areas (1σ) for Italian unstressed final vowels.

Overlap of areas and centralization

As is widely known, the ellipses shown in Figures 1, 2 and 3 define a portion of the f1/f2 plane which includes about 67% of the vowels belonging to the given category: in other words this kind of representation gives an immediate idea of the most probable values for f1 and f2 for each vowel type. Observing he data distribution more closely, the following phenomena become clear:

1) within each diagram, data coming from adjacent vowel categories are strongly overlapped; as an example of this statement Fig. 4 shows the same data as Fig. 2 but the ellipse axes are now ±2σ, corresponding to the delimitation of a portion of the f1/f2 plane containing about 95% of the data of each vowel type. In this graph the overlaps between the ellipses are dramatic. Similar results would be obtained applying the same procedure to the data of Figures 1 and 3.

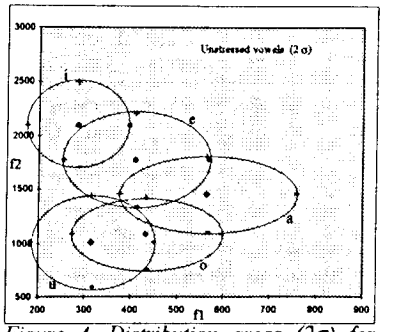

Figure 4. Distribution areas (2σ) for Italian unstressed non final vowels.

2) the comparison of mean values for similar vowel categories in different stress conditions indicates a tendency to the centralization of all vowels when unstressed. In Figure 5 the mean values of stressed vowels are compared those obtained in the unstressed condition. As in Italian the stressed vowel couples [e, ɛ] and [ɔ, o] are reduced respectively to [e] and [o] when unstressed, we compared the unstressed data with the average between the two corresponding stressed vowels. Each arrow-ended line in the graph corresponds to the 'direction' of this centralization. These lines tend to meet in a very small area. The little square in the middle of Figure 5 indicates the center point of this area.

The ideal ellipse defining such area has an f1 = 332±24 Hz and an f2 = 1350±46 Hz.

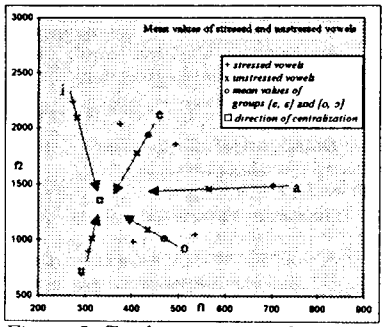

Figure 5. Tendency to centralization in unstressed vowels.

Many authors [7], [8], [9] have stated the importance of another kind of ideal point, named centroid and defined as "*...the grand mean of all measured formant frequency of the vowel system per speaker.*"[9:159]. The centroid is generally calculated either for only stressed or for only unstressed vowels, and gives an idea of a tendency that is internal to that vowel system. Our procedure, on the contrary, describes a tendency related to the different behaviour of the speakers when uttering a stressed vowel and an unstressed one. We calculated the coordinates of the centroid (average for all twenty speakers) of our stressed and unstressed vowels in order to compare them with the ideal point (from now on CT 'Centralization Tendency') calculated with our method. The results are:

centroid:		
stressed	f1 = 441	f2=1505
unstressed	f1 = 401	f2=1486
CT :	f1=332	f2=1350

CT seems to indicate a point in the f1/f2 plane quite different from the centroids. The ideal vowel corresponding to CT is higher and more velar (back) than the centroids.

Vowel duration

As expected, stressed vowels show a much longer duration (averagely about 97 ms) than final (62 ms) and non final unstressed (57 ms) ones (see Fig.6). It's very interesting to observe, on the other hand, that all groups of vowels show a very regular correlation between duration and openness. This correlation seems to be particularly relevant for stressed vowels, which range from a minimum of 84±29 ms for [i] and 86±29 ms for [u], to a maximum of 119±38 ms for [a].

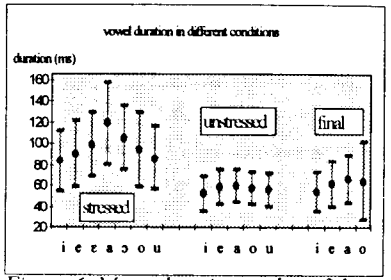

Figure 6. Mean duration and confidence interval of Italian vowels.

A comparison between the duration of stressed vowels in open syllables with vowels in closed syllables is shown in Figure 7. Unlike what is generally accepted in literature [2], [4], [10], [11] our data show that no significant difference exists between the two groups.

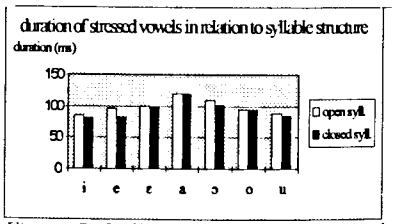

Figure 7. Differences of stressed vowel duration between open and closed sillables.

COMMENT

Comparing our stressed vowel data to the unstressed ones a clear tendency to centralization is observable. We have tried to express this tendency in terms of coordinates of a particular ideal point as shown in Fig. 5. As a consequence, our method seems to state the loss of the symmetry introduced with the concept of centroid.

Moreover, centralization is classically used in synchronic and diachronic phonetics for phenomena of vowel alternation or change resulting into a *schwa* [ə] (corresponding approximately to f1=500, f2=1500 Hz). In our case, on the contrary, we use the term centralization to indicate a tendency, i.e. a process of partial convergence towards an ideal point in the f1/f2 plane; which does not coincide with *schwa* (in this sense we use the term "point" instead of "vowel" in order to emphasize the difference).

The vowel [ə] has a well distinct nature and has also a pertinent phonological value in many languages. Presently Italian seems not to belong to the class of the languages having a [ə] in their vocalic system. Though, on the base of the results herein presented, a class of partially centralized vowels should be introduced for a more complete description of unstressed (non final and final) vocalic sounds.

In this paper many results seem to differ (more or less slightly) from the normally accepted description of Italian vowel system. It is our opinion that this is mainly due to the choice of using "spontaneous" speech materials instead of laboratory speech.

Many further investigations need to be carried out on our data, some of the further planned developments are:

-analyses of female data and description of the differences between the two groups;

-effects of coarticulation with adjacent consonants and of assimilation with vowels in adjacent syllables (some relations between these effects and the entity of the overlaps of vocalic areas are foreseen);

-investigations on the role of other measured acoustical parameters such as f_0, f3, maximum energy within the vowel;

-diatopic analyses of data, namely: a comparison between the vowel systems in different regional standards of Italian.

REFERENCES

[1] Ferrero, F. (1968), "Diagrammi di esistenza delle vocali italiane", *Alta Frequenza*, vol. 37, pp. 9-31.

[2] Fava E., Magno Caldognetto E. (1976), "Studio sperimentale delle caratteristiche elettroacustiche delle vocali toniche e atone in bisillabi italiani", in R. Simone et al. (eds), *Studi di fonetica e fonologia*, Roma, Bulzoni 1976, pp.35-79.

[3] Ferrero F., Magno Caldognetto E., Vagges K., Lavagnoli C. (1978), "Some acoustic characteristics of the Italian vowels", *Journal of Italian linguistics* 3, pp.87-96.

[4] Marotta G. (1984), "*Aspetti della struttura ritmico-temporale in italiano. Studi sulla durata vocalica*", Pisa, ETS.

[5] Albano Leoni F., Maturi P., (1994), "Didattica della fonetica italiana e parlato spontaneo", in Giacalone Ramat A., Vedovelli M., (eds) *Atti del XXVI Convegno nazionale SLI*, Roma, Bulzoni, pp.153-164.

[6] Albano Leoni F., Caputo M.R., Cerrato L., Cutugno F., Maturi P., Savy R., (1994), "Il vocalismo dell'italiano. Analisi di un campione televisivo", in Perrone B., (ed) *Atti del XXII Convegno nazionale AIA*, Lecce, pp.419-424.

[7] Liljencrants J., Lindblom B., "Numerical simulation of vowel quality systems: the role of perceptual contrast", *Language*, Vol.48, 1972, pp.839-862.

[8] Disner S.F., (1980), "Insights in vowel spacing: results of a language survey.", *UCLA Working papers in phonetics*, Vol.50, pp.70-92.

[9] Koopmans-van Beinum F.J., (1983), "Systematics in vowel systems", in van den Broecke M., van Heuven V., Zonneveld W., (eds) *Sound Structures*, FORIS Publications, Dordrecht, pp.159-171.

[10] Salza P.L., Sandri S. (1987), "Consonant-to-vowel durational effects in Italian", *CSELT Technical Report*, Vol. XV n.1, February, pp.61-66.

[11] Farnetani E., Kori S. (1984), "Effects of syllable and word structure on segmental duration in spoken Italian", *Quaderni del Centro di studio per le ricerche di fonetica*, 3 , pp.143-188.

A Microsoft Excel 5.0 version of our database may be accessed via anonynous ftp on Internet. The address of ftp server is: dsna1.na.infn.it. Go to the directory CIRASS and use binary mode to retrieve the file named vocali.xls.

Acknowledgements This work was realized with the help of all CIRASS members for measurements and data elaboration. Moreover we wish to thank Loredana Cerrato and Pietro Maturi that took part to all the preceeding phases of the project.