A DYNAMIC APPROACH OF VOWELS SYSTEMS IN ITALIAN

D. Poch-Olivé
Autonomous University of Barcelona, Barcelona, Spain

B. Harmegnies
University of Mons-Hainaut, Mons, Belgium

ABSTRACT

A native speaker of Italian was first recorded on the occasion of a natural conversation. After transcription of the whole corpus, 210 words, each containing one token of a given vowel, were selected. They were read in laboratory conditions by the speaker. A preliminary statistical treatment, focused both on the [e]/[ɛ] and the [o]/[ɔ] differences, confirms that whereas tendencies to merge those categories exist in Italy, the recorded informant does use a 7-vowel system. The formants frequencies are further processed both by means of the δ index and of discriminant analysis. The statistical treatment confirms the existence of phenomena similar to those observed in other languages of the Romance group, from both qualitative and quantitative points of view.

INTRODUCTION

Harmegnies and Poch have been carrying out a joint study project focused on the dynamics of vowels systems under the effect of speaking style. Their experiments have dealt with languages of the Romance group, i.e., Castilian Spanish [1], Catalan [2], Belgian French [3] and European Portuguese [4]. Languages in this group are very interesting, because whereas they derive from the same origin, they are characterized by vowel systems differing one from another in quite a large variety of ways, i.e., 1. by the number of units in the system (from 5: Spanish, up to 14: French); 2. by the presence of phonological reduction in the system (Catalan and Portuguese) or its absence (Spanish and French); 3. by the existence of central vowel(s) in several of them (Catalan, French and Portuguese) although Spanish has no such vowel.

Laboratory speech (words lists reading) and spontaneous speech (spontaneous conversation), at least, have been considered in all those languages. When compared to laboratory speech, spontaneous speech may be characterized by 1.schwa-like tendency (not in all the languages); 2.reinforced timbres variability (not in all the languages); 3.increased overlapping of the vowels clusters (in all the languages).

Further research requires to use other languages both in order to evaluate the universality of the phenomena observed, and to seek a wider understanding of them.

Particularly, among the languages involved up to now, Spanish is the only one without central vowel; the study of another language characterized by the same feature therefore appears quite desirable.

In this paper, we apply the previously developped methodology to Italian, which possesses a seven-unit system, without central vowel, and free of phonological reduction.

EXPERIMENTAL SETTING

The *spontaneous speech* sample was drawn from a natural conversation held with a male Italian speaker. This was born and lived in Napoli, and spoke standard Italian. The talk lasted about one hour and took the form of a semi-directive interview led by an experimenter, where the subject was suggested to evoke various themes, such as Italian food, his birth place, or current occupation.

On the basis of an exhaustive transcription of the whole recording, a set of 210 vowels (i.e., 30 samples of each of the 7 Italian vowels) was randomly selected. Each word containing a selected vowel was randomly put in a list that the speaker was afterwards asked to read, in order to produce the *laboratory speech* sample.

RESULTS

The sampled vowels

Finally, 420 vowels, organized in pairs with invariant status in spontaneous- and lab speech, were retained. Their first and second formants frequencies were measured both in spontaneous- and laboratory speech by means of a DSP 5500 KAY analyzer, which provided a 20 Hz resolution on the whole frequency span. These frequency values are summarized in tables 1 and 2.

Table 1. Averages (m) and standard deviations (s) of the first formants frequencies in laboratory- ("Lab") and spontaneous ("Spt") speech.

	Lab		Spt	
	m	s	m	s
[i]	311.23	37.37	391.50	62.58
[e]	396.40	34.21	439.23	51.83
[ɛ]	442.50	33.14	505.76	34.20
[a]	591.26	65.52	595.33	69.71
[ɔ]	459.50	21.10	510.13	38.21
[o]	429.30	40.17	457.60	59.08
[u]	347.46	44.05	394.73	52.40

Table 2. Averages (m) and standard deviations (s) of the second formants frequencies in laboratory- ("Lab") and spontaneous ("Spt") speech.

	Lab		Spt	
	m	s	m	s
[i]	1951.53	71.17	1807.00	94.15
[e]	1697.60	100.66	1651.30	99.18
[ɛ]	1717.33	73.78	1603.13	96.22
[a]	1305.66	109.45	1326.70	104.88
[ɔ]	931.36	75.33	1092.03	158.00
[o]	988.86	150.47	1103.96	140.23
[u]	916.43	180.47	980.16	179.32

The vowel system

Our procedure of vowel selection was based upon the idea that the Italian vowel system is composed of the 7 open vowels [i], [e], [ɛ], [a], [ɔ], [o] and [u]. Nevertheless, as speakers of Italian tend nowadays to merge the realizations of /e/ and /ɛ/, as well as the ones of /ɔ/ and /o/, we tried to determine whether the realizations of the corresponding vowels by our informant were to be considered as belonging to different categories.

The 30 presumed [e] were compared to the 30 presumed [ɛ] in both styles. The comparison revealed significant differences along the first formant axis (laboratory speech: Mann-Whitney's U = 151, p<.0001; spontaneous speech: Mann-Whitney's U = 126, p<.0001), but not along the second formant axis (laboratory speech: Mann-Whitney's U = 405, p=.5; spontaneous speech: Mann-Whitney's U = 333, p=.0845).

The 30 presumed [ɔ] were, in turn, compared to the 30 presumed [o]. The comparison revealed significant differences along the first formant axis, as well, (laboratory speech: Mann-Whitney's U = 232, p=.0012; spontaneous speech: Mann-Whitney's U = 205, p=.0003), but not along the second formant axis (laboratory speech: Mann-Whitney's U = 395, p=.4155; spontaneous speech: Mann-Whitney's U = 394, p=.4113).

In other words, it seems reasonable to consider that our informant uses a 7-vowel system, since the /e/-/ɛ/ and the /ɔ/-/o/ pairs appear as significantly differenciated, at least relative to the aperture dimension.

The dynamics of the system

Previous research [1-4] has showed that vowels tend to centralize, i.e., to appear closer to the F1/F2 plane center, when uttered in spontaneous speech. In order to test this hypothesis, we computed, for each inter style pair of vowel, a

centralization index, δ [1]. This one is defined as the difference between the Euclidean distance from the observed vowel to schwa in laboratory speech and the Euclidean distance from the observed vowel to schwa in spontaneous speech, i.e., the amount of displacement towards schwa:

$$ED_{Lab}=\sqrt{(F1-500)^2+(F2-1500)^2} \quad (1)$$

$$ED_{Spt}=\sqrt{(f1-500)^2+(f2-1500)^2} \quad (2)$$

$$\delta=ED_{Lab}-ED_{Spt} \quad (3)$$

where **ED** stands for Euclidean Distance, **F** symbolizes formants values for the laboratory speech sample, and **f** for the spontaneous speech sample. As can be predicted from equation 3, positive values only of δ denote centralization, the magnitude of which is measured by δ. Moreover, as δ is a difference index, its significance can be tested by means of paired two-sample inferential procedures; in this case, the null hypothesis is that the difference between ED_{Lab} and ED_{Spt} does not significantly differ from zero.

Values of the δ index are presented in table 3, together with the results of inferential tests. The paired Student t test has been used, as well as the Wilcoxon matched pairs T test. As can be seen from table 3, centralization turns out to be significant for all vowels, but [a]. It is to be noticed that since both the parametic and the non-parametric procedures deliver the same conclusions, possible artifacts caused by the shapes of the δ distributions should be considered very unlikely. Also, the results of the ranked-based procedure would be held constant under any monotonic transform of the frequencies (such as mel transform, e.g.).

The centralizing tendency revealed by our treatment goes together with increase of the surface occupied by the vowels in the F1/F2 plane, in spontaneous speech, relative to laboratory speech. This can be observed from the relationships between the formant variabilities. As shown in tables 1 and 2, the standard deviations of the first formants values are systematically greater in spontaneous speech than in laboratory speech. The tendency is less obvious in F2, where the standard deviations are less different one from another.

Table 3. δ values, student paired t-test statistic ("t") with probability under the null hypothesis ("p"), and Wilcoxon matched pairs test normal approximation ("z_w") with probability under the null hypothesis ("p").

	δ	t	p	z_w	p
[i]	158.6885	8.58	<.001	4.72	<.0001
[e]	47.0357	2.99	.006	2.56	.0104
[ɛ]	102.6575	7.45	<.001	4.65	<.0001
[a]	11.5071	.68	.499	.54	.5857
[ɔ]	152.3713	6.52	<.001	4.62	<.0001
[o]	115.7178	4.29	<.001	3.43	.0006
[u]	72.1478	2.80	.009	2.48	.0132

The combination of the decreased formants differences caused by centralization, together with the increased formant variability within each vocalic category decreases differenciations in the whole F1/F2 spontaneous speech system: not only are the vocalic clusters closer one to another in the F1/F2 space, they are moreover less homogeneous. The spontaneous speech system therefore seems to have reached a more pronounced degree of disorganization than the laboratory speech one.

As a general rule, increased entropy in any system implies that the system is less informant: one may therefore expect the lab speech F1/F2 system to convey more information than the spontaneous speech one. Recognizing the elements of the system should therefore be more a hazardous task in the spontaneous- than in the lab speech sample.

In order to test this hypothesis, we performed 2 discriminant analyses (one in spontaneous speech and one in lab speech), with the vowels as *a priori* categories and the formant values as discriminant variables. Once computed the discriminant functions, a recognition task was simulated in each subsample. Their results are presented in tables 4 and 5.

Table 4. Confusion matrix from the simulated vowel recognition task in laboratory speech. Actual groups are in rows, and predicted group membership in columns.

	/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
/i/	*93.3*	3.3	3.3	0	0	0	0
/e/	6.7	*70.0*	23.3	0	0	0	0
/ɛ/	0	16.7	*83.3*	0	0	0	0
/a/	0	0	3.3	*90.0*	0	6.7	0
/ɔ/	0	0	0	0	*86.7*	13.3	0
/o/	0	0	0	0	40.0	*46.7*	13.3
/u/	0	0	0	0	3.3	6.7	*90.0*

Table 5. Confusion matrix from the simulated vowel recognition task in spontaneous speech. Actual groups are in rows, and predicted group membership in columns.

	/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
/i/	*76.7*	16.7	6.7	0	0	0	0
/e/	16.7	*60.0*	23.3	0	0	0	0
/ɛ/	0	10.0	*86.7*	0	0	3.3	0
/a/	0	0	6.7	*80.0*	6.7	6.7	0
/ɔ/	0	0	3.3	6.7	*73.3*	13.3	3.3
/o/	0	0	3.3	3.3	33.3	*33.3*	26.7
/u/	0	3.3	0	0	3.3	20.0	*73.3*

The recognition procedure clearly appears safer when performed on the basis of laboratory speech samples. The overall correct recognition is, in this case, 80%, although it decreases to 69% in spontaneous speech. This observation thus confirms that vowels in the F1/F2 plane are more differenciated in laboratory- than in spontaneous speech.

CONCLUSIONS

This study confirms, for Italian, tendencies already pointed out for other languages of the Romance group. Vowels in spontaneous speech are realized closer to the F1/F2 plane center; they moreover are better differenciated one from another in laboratory- than in spontaneous speech. Although drawn from a single-speaker experiment, those findings constitute an interresting account to the study of the universality of the reported phenomena. Further research should nevertheless seek to confirm the findings, and to refer observed variabilities to the ones caused by interindividual differences and sociolinguistic factors.

ACKNOWLEDGMENT

The recordings have been collected by David Puigvi, at the CIRASS (University of Napoli), under the supervision of Dr P. Maturi. Transcriptions have been performed by M. Piccaluga.

REFERENCES

[1] HARMEGNIES, B., POCH-OLIVE, D., "A study of style-induced vowel variability: laboratory versus spontaneous speech in Spanish", *Speech Communication*, 1992, 11, 429-437.

[2] BLECUA, B., POCH-OLIVE, D., HARMEGNIES, B., "Variaciones en la organizacion de los sistemas vocalicos del espanol y del catalan en funcion del estilo de habla", *Actas de las Jornadas Internacionales de Linguistica Aplicada*, Universidad de Granada, 1993, 22, 19-31.

[3] POCH-OLIVE, D., HARMEGNIES, B., "Variations structurelles des systèmes vocaliques en français et espagnol sous l'effet du style de parole", *Journal de Physique*, 1992, 4, 283-286.

[4] DELPLANCQ, V., HARMEGNIES, B., POCH-OLIVE, D., "Effets du style de parole sur la réduction vocalique en portugais", to be printed in *Verbum*.