VISIBLE ARTICULATORY CHARACTERISTICS OF THE ITALIAN STRESSED AND UNSTRESSED VOWELS

Emanuela Magno Caldognetto, Kyriaki Vagges, Claudio Zmarich
Centro di Studio per le Ricerche di Fonetica, CNR
Via Anghinoni, 10 35121 Padova, Italy

ABSTRACT

This research focuses on the study of the multidimensionality of the visible articulatory movements in the production of the Italian stressed and unstressed vowels. Lip and jaw movements were recorded and analysed with a fully automatic real-time system for 3D kinematics data acquisition. The data obtained show that *jaw opening* and *lower lip protrusion* were the most relevant articulatory parameters in distinguishing among the vowels.

INTRODUCTION

Integration of articulatory data with acoustic and perceptual data is fundamental in developing a phonetic theory of vowels and can find important applications in linguistic description [1,2] (cross-language comparison, articulatory features system), psycholinguistic research [3] (lip reading, bimodal perception) or technological applications [4] (audio-visual speech synthesis and recognition systems).

The Italian vowel system is considered to be very simple [1] since it may be described in terms of the features *high / low* and *front / back*. The feature *rounded* cooccurs with the feature *back* and *protrusion* always cooccurs with *rounding*.

The aim of this research is to define the visible articulatory parameters in the production of the Italian stressed and unstressed vowels, to individuate the cooccurrence of various parameters and finally to evaluate the relation of these parameters to the phonetic linguistic features.

METHOD

Lip and jaw movements were recorded and analysed with ELITE [5], a fully automatic, real-time system for 3D kinematics data acquisition. It uses small, non obtrusive, passive markers of 2mm of diameter, realised by reflective paper, attached onto the speaking subject's face. In this study the markers were placed on the central points of the vermilion border of the upper lip and of the lower lip, at the corners of the lips, and at the centre of the chin. The markers placed on the tip of the nose and on the lobes of the ears served as reference points to eliminate the effects of the head movement. The following articulatory parameters corresponding to phonologically significant features were analysed:

- *lip height* (LH), calculated as the distance between the markers placed on the central points of the upper and lower lips; this parameter may be correlated with the feature *high/low*.
- *lip width* (LW), corresponding to the distance between the markers placed at the corners of the lips, a parameter which correlates with the feature *rounded / unrounded*.
- *jaw opening* (JO), corresponding to the distance between the markers placed at the centre of the chin and the tip of the nose. This distance is primarily due to the jaw opening but it is also influenced by the movement of the skin of the chin. This parameter is correlated with the feature *high / low*.
- *anterior - posterior movement* of the *upper lip* (UP) and *lower lip* (LP), calculated as the distance between the markers placed on the central points of either the upper or lower lip and the line passing from the markers placed on the lobes of the ears. This parameter correlates with the feature *protruded / retracted*.

The visible articulatory movements of 6 subjects (4 females and 2 males), speakers of northern Italian, were recorded and analysed. All the subjects were university students, aged between 19 and 22 and were paid volunteers. They repeated five times, in random order, each of the 7 stressed /a/, /ɛ/, /e/, /i/, /ɔ/, /o/, /u/ and the 5 unstressed, /a/, /e/, /i/, /o/, /u/, Italian vowels. The vowels were in the first syllable of disyllabic (/'tasti/, /'tɛsi/, /'tesi/, /'tisi/, /'tɔsko/, /'tosko/, /'gusto/) or trisyllabic (/ta'stare/, /te'stare/, /ti'sane/, /to'skane/, /gu'stare/) words, and were preceded by a /t/ and followed by /s/ (with one exception, i.e. /gu'stare/). They occured within the carrier phrase "dico _____ chiaramente" (I say _____ clearly).

A synchronous recording of the acoustic signal was also obtained.

Portions of the articulatory signal corresponding to the vowel to be analysed were segmented on the basis of the acoustic speech signal. Since the dynamic aspects of the articulatory parameters were not taken into consideration in this study, a single point characterising the vowel was individuated for each articulatory parameter. The data were normalised subtracting the values related to the position of the lips and jaw at rest, from each parameter obtained, for each vowel and each subject. This assured the comparability of the results independently of the subjects variability in the shape and size of the articulators. The so obtained data correspond to the real extension of the lip and jaw movements and may also be connected to data relating to the internal borders of the lips [4].

RESULTS

The analysis of the normalised data showed the mutual relations between all the parameters examined: a clear correlation between the *upper* and the *lower lip protrusion* for both stressed and unstressed vowels, (r=.82 and r=.83 respectively), and a negative correlation between *lip width* and *upper lip protrusion* (r=-.81 and r=-.83), as well as between *lip width* and *lower lip protrusion* (r=-.73 and r=-.80). Stressed vowels showed also a correlation between *lip height* and *jaw opening* (r=.85). For both stressed and unstressed vowels there was no significant correlation between *lip height* and *lip width* and between *lip height* and *upper* or *lower lip protrusion*. The presence or absence of correlation observed in this study is congruous with previous results reported for English [6,7] and French [8] independently of the instrumentation or reperee points used in defining the parameters.

Stressed and unstressed vowels were analysed with two-way ANOVAs (7 or 5 vowels respectively and 6 subjects as a between factor) to assess their effect on each of the articulatory parameters examined. Post hoc Tukey multiple comparisons were carried out when the vowel effects were significant. Only the data significant at $p<.01$ will be discussed. The normalised mean values, pooled over the 6 subjects and the 5 repetitions for each parameter and each vowel are reported in Table 1. The values may be either positive or negative depending on the parameter taken into consideration. For example, LW values are negative when the distance between the corners of the lips decreases with respect to their distance at rest, as is evident for both stressed and unstressed /u/, while positive LW values correspond to an increased distance with respect to the values at rest, as is the case of the unrounded vowel /i/ in both stressed and unstressed position. UP and LP may also show both positive and negative normalised mean

Table 1. Normalised mean values (mm) pooled over subjects and repetitions for each articulatory parameter and each vowel.

		/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
LH	stressed	8.6	9.6	13.6	15.0	15.1	8.7	7.7
	unstressed	7.6	8.7		10.6		8.1	7.1
JO	stressed	6.6	7.8	12.6	14.1	11.2	3.3	1.8
	unstressed	5.3	6.7		9.0		2.6	2.0
LW	stressed	0.1	0.5	1.1	0.9	-5.1	-5.1	-6.1
	unstressed	1.2	0.8		1.2		-3.3	-5.3
UP	stressed	-1.1	-0.3	-2.0	-2.1	2.6	3.9	4.4
	unstressed	-1.2	-1.0		-1.5		3.2	3.9
LP	stressed	-1.4	-1.2	-3.3	-2.9	0.9	2.8	3.6
	unstressed	0.6	-1.1		-1.9		2.2	3.4

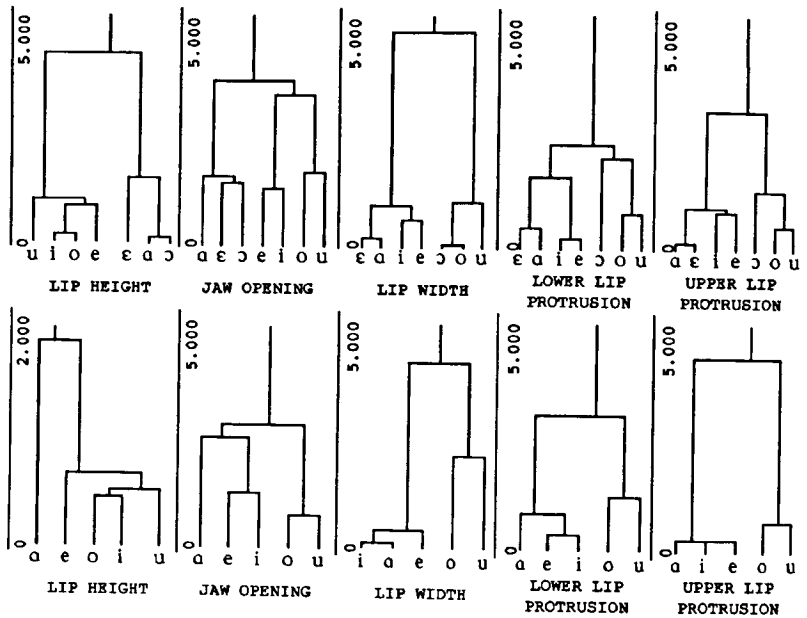

Figure 1. Hierarchical clustering of the stressed and unstressed vowels with respect to the five articulatory parameters.

values, while LH and JO are always positive.

The results of the ANOVA showed that *jaw opening* is the articulatory parameter that better distinguishes both stressed and unstressed vowels since it defines 4 degrees of jaw opening. In fact, Fig. 1 shows that stressed vowels are clustered in 4 groups, /u,o/, /i,e/, /ɛ,ɔ/, and /a/. As for the unstressed vowels, in Fig. 1, JO distinguishes /a/, /e/, /i/ and /u,o/.

LH, which is traditionally considered to be parallel to *jaw opening*, does not identify all the degrees of opening defined by JO (see Fig. 1). Moreover, the extension of its movement always shows greater values than JO, cf. Table 1. It is clear that lips not only move in synergy with the jaw, but also in an independent specific manner.

LW divides both stressed and unstressed vowels in two groups: rounded vowels and unrounded vowels, see Fig. 1.

As for the two protrusions, LP is the parameter that best distinguishes both stressed and unstressed vowels. As shown in Fig. 1, stressed and unstressed vowels are divided into 4 groups, i.e. two degrees of protrusion and two degrees of retraction. In particular, for stressed vowels, a higher degree of protrusion characterises /u/ and /o/ with respect to /ɔ/, while /a/ and /ɛ/ are more retracted than /i/ and /e/, see Table 1.

Using the parameters resulting most significant for distinguishing the vowels (*jaw opening, lower lip protrusion* and *lip width*), a three dimensional representation of the stressed and unstressed vowel space was plotted in Figs 2a and 2b respectively. As can be observed there is a tendency to reduce the values of the parameters from the stressed to the unstressed condition, even though the trend is not systematic.

DISCUSSION

Our data confirm the cooccurrence of rounding and protrusion for the Italian language. In fact, all the vowels with positive values of *lip width*, i.e. /i,e,ɛ,a/, also have negative values for both *upper*, and *lower lip protrusion*. That is, unrounded

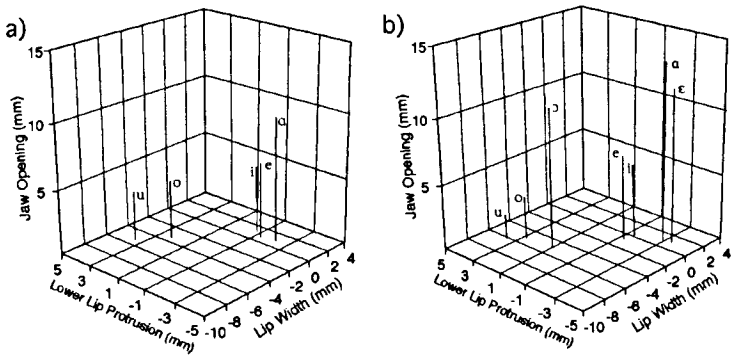

Figure 2. 3D representation of the stressed and unstressed vowel space.

vowels are always also non protruded. Similarly, vowels with negative *lip width* values, i.e. the rounded vowels /ɔ,o,u/, are characterised by positive values of *upper* and *lower lip protrusion*, that is, they are also protruded.

Jaw opening and *lower lip protrusion* are the parameters that better distinguish the vowels. It should be noted though, that differences in *jaw opening* with respect to *lip height* may be due to the marker placed on the chin: the position of this marker was influenced not only by the jaw opening but also by the movement of the skin especially during the lip protrusion.

Based on the values of the parameters analysed, the reduction of the unstressed with respect to the stressed vowels was confirmed. Moreover, the unstressed mid vowels are more similar to the stressed mid-high /e/ and /o/ rather than to the mid-low stressed /ɛ/ and /ɔ/.

REFERENCES

[1] Ladefoged P., & Maddieson I., (1990) *Vowels of the world's languages*, Journal of Phonetics 18, 93-122.

[2] Fischer-Jørgensen E., (1985) *Some basic vowel features, their articulatory correlates, and their explanatory power in phonology*, in Fromkin V.A. (Ed), Phonetic Linguistics. Essays in Honor of Peter Ladefoged, Academic Press, 79-99.

[3] Summerfield Q., (1987) *Some preliminaries to a comprehensive account of audio-visual speech perception*, in Dodd B. & Campbell R. (Eds.), Hearing by Eye: The Psychology of Lip-Reading, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 3-51.

[4] Benoît C., Lallouache T., Mohamadi T., & Abry C., (1992) *A set of French visemes for visual speech synthesis*, in Bailly G., Benoît C., & Sawallis T.R. (Eds.), Talking machines: Theories, Models, and Designs, North-Holland, Amsterdam, 485-504.

[5] Magno Caldognetto E., Vagges K., Pedotti A., Ferrigno G., (1994) *Parametri articolatori labiali e mandibolari nelle vocali cardinali dell' Italiano*, Atti delle III Giornate di Studio del G.F.S. :Le vocali, Padova 1992, 75-85.

[[6] Fromkin V., (1964) *Lip positions in American English vowels*, Language and Speech 7, 217-225.

[7] Linker W., (1982) *Articulatory and acoustic correlates of labial activity in vowels: A cross-linguistic survey*, UCLA, Working Papers in Phonetics 56, 1-134.

[8] Abry C., & Boe L.J., (1986) *"Laws" for lips*, Speech Communication 5, 97-104.