Coproduction of Speech and Emotions: Visual and Acoustic Modifications of Some Phonetic Labial Targets

Emanuela Magno Caldognetto, Piero Cosi, Carlo Drioli, Graziano Tisato, Federica Cavicchio

Laboratory of Phonetics and Dialectology, ISTC-CNR
Institute of Cognitive Sciences and Technology, ITALY
e-mail: {magno,cosi,drioli,tisato,cavicchio}@csrf.pd.cnr.it

Abstract

This paper concerns the bimodal transmission of emotive speech and describes how the expression of joy, surprise, sadness, disgust, anger, and fear, leads to visual and acoustic target modifications in some Italian phonemes. Current knowledge on the audio-visual transmission of emotive speech traditionally concerns global prosodic and intonational characteristics of speech and facial configurations. In this research we intend to integrate this approach with the analysis of the interaction between labial configurations, peculiar to each emotion, and the articulatory lip movements defined by phonetic-phonological rules, specific to the vowels and consonants /'a/, /b/, /v/ ([1], [2]). Moreover, we present the correlations between articulatory data and the spectral features of the co-produced acoustic signal[1].

1.Introduction

Traditionally, facial and speech acoustic cues (both segmental and supra-segmental) conveying emotions have been studied separately (from the first studies on facial configurations [3] and on acoustic characteristics [4], [5] to more recent reviews of the state of the art in these fields [6], [7], [8]). The relevance of the interaction between audio and visual modalities in the transfer of emotions has been stressed from the perceptual point of view ([9], [10], [11], [12]). More effective results in the description of emotive speech could be achievable studying the labial movements, which are the effect of the compliance with both the phonetic-phonological constraints and the configurations required for the encoding of emotions. To be completely satisfactory, the studies on bimodal audiovisual emotive speech should explain the complex interactions among four categories of indices: 1) the visual cues related to the lips, the jaw, and the tongue articulatory movements, carrying the information on visemes ([1], [2]); 2) the labial and facial visual cues related to the emotions; 3) the acoustic cues conveying linguistic segmental, supra-segmental, prosodic, and intonational, characteristics of speech; 4) the acoustic correlates of emotions: F0 parameters, voice quality, prosody, and intonation. Following this approach, we focused this research on the quantification of the labial articulatory parameters modifications induced by the different emotions, and on the definition of their acoustic correlates (see the recent papers on smile, [13], [14]). The obtained results will be useful for the formulation of cognitive theories on coproduction and perception of linguistic and paralinguistic information, mixed in emotive speech, and for the working out of speech technologies, i.e. bimodal speech synthesis (Talking Heads) and recognition systems.

2. Method

In order to collect the articulatory and acoustic data, an automatic optotracking movement analyzer for 3D kinematics data acquisition (ELITE) was used, which also allows a synchronous recording of the acoustic signal (for previous applications of this data acquisition system to the definition of Italian visemes on an articulatory basis, see [1], [2]). This system ensures high accuracy (100 Hz sampling rate, maximal error of 0.1 mm for a 28x28x28 cm cube) and minimum discomfort to the subject because it tracks the infrared light reflected by small (2 mm diameter), passive markers glued on different points of the external lips contour and of the face, following the scheme in Fig.1. Only the data relative to lip markers are presented in this work. A male University student, who speaks a northern regional Italian and with recitation skills, pronounced two phonological structures 'VCV, corresponding to two feminine proper names: "Aba" /'aba/ and "Ava" /'ava/, simulating, on the basis of appropriate scenarios, six emotional states: anger (A), joy (J), fear (F), sadness (SA), disgust (D) and surprise (SU), apart from the neutral one (N), corresponding to a declarative sentence. This 14 words set was repeated many times in random

[1] Part of this work has been sponsored by COMMEDIA (COMunicazione Multimodale di Emozioni e Discorso in Italiano con Agente animato virtuale, CNR Project C00AA71), PF-STAR (Preparing Future multiSensorial inTerAction Research, European Project IST- 2001-37599, http://pfstar.itc.it) and TICCA (Tecnologie cognitive per l'Interazione e la Cooperazione Con Agenti artificiali, joint "CNR - Provincia Autonoma Trentina" Project).

order, leading to a total of 107 recordings[2]. For each acquisition session the articulatory data at resting position has been recorded as well, and the extracted parameters have been normalized with respect to this values.

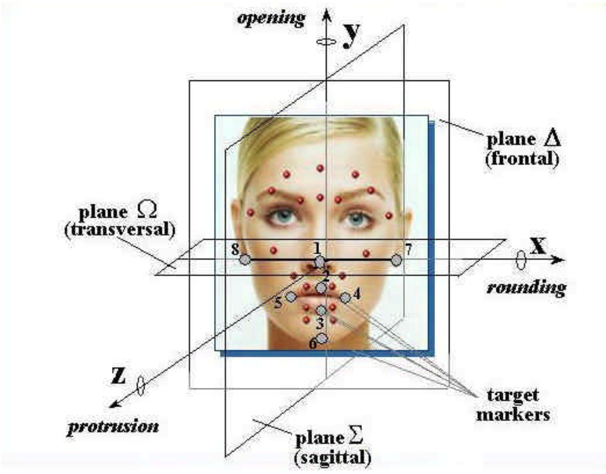


Figure 1: *Position of the 28 reflecting markers and of the reference planes for the articulatory movement data collection.*

Concerning the phonological structure of the selected words, we chose the labial voiced stop /b/ and the labiodental voiced fricative /v/, as they belong to the 2 classes of consonants for which the lips are the active articulators and execute phonologically specified movements. In the Italian vowel system, the central open vowel /a/ offers the widest lip opening and is characterized by the largest articulatory space [2]. The parameters selected to quantify the labial configuration modifications are the following:

• Lip Opening (LO), calculated as the distance between markers placed on the central points of the upper and lower lip vermillion borders ; this parameter correlates with the HIGH-LOW phonetic dimension.

• Lip Rounding (LR), corresponding to the distance between the left and right corners of the lips, which correlates with the ROUNDED-UNROUNDED phonetic dimension: negative values correspond to the lip spreading.

• Anterior/posterior movements (Protrusion) of Upper Lip and Lower Lip (ULP and LLP), calculated as the distance between the marker placed on the central points of either the upper and lower lip and the frontal plane containing the line crossing the markers placed on the lobes of the ears and perpendicular to Ω plane. These parameters correlate with the feature PROTRUDED-RETRACTED: negative values quantify the lip retraction.

• Upper and Lower Lip vertical displacements (UL, LL), calculated as a distance between the markers placed on the central point of either upper and lower lip and the transversal plane Ω, passing through the tip of the nose and the markers on the ear lobes. Hence, positive values correspond to a reduction of the displacement of the markers from the Ω plane. As told before, these parameters are normalized in relation to the lip resting position.

• Left and Right Corner horizontal displacements (LCX and RCX), calculated as the distance between the markers placed on the left and the right lip corner and the sagittal Σ plane passing through the tip of the nose and perpendicular to the Ω plane (these parameters are not visualized in Fig. 2).

• Left and Right Corner vertical displacements (LCY and RCY), calculated as the distance between the markers placed on the left and right lip corner and the transversal plane Ω, containing the line crossing the markers placed on the lobes of the ears and on the nose (these parameters are not visualized in Fig. 2).

• The asymmetry parameters (ASYMX and ASYMY) were calculated as the difference between right and left corner position along the x (RCX-LCX) and y (RCY-LCY) axes. Both for ASYMX and ASYMY values different from zero indicate the presence of an asymmetry. Positive values for ASYMY mean that the right lip corner moves in an asymmetric higher position along the vertical axis than the left corner. Positive values for ASYMX indicate that the lips are displaced in a right asymmetrical way along the horizontal axis.

Fig. 2 proposes some of the labial kinematic parameters considered in this study and relative to /’aba/ expressing disgust: LO, LR, ULP, LLP, UL, LL, ASYMX and ASYMY.

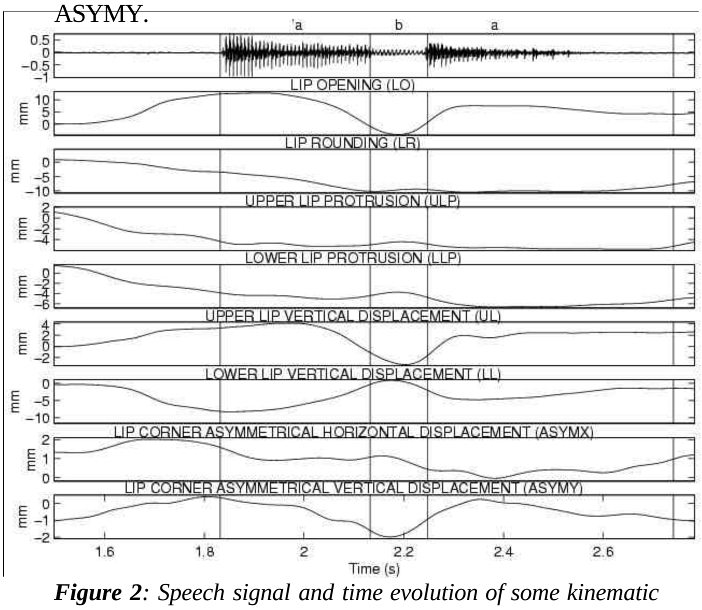


Figure 2: *Speech signal and time evolution of some kinematic parameters associated with the sequence /’aba/ expressing disgust.*

[2] Due to the complexity of the movement tracking algorithm, some acquisitions resulted unreliable. For this reason, 7 cases were used for A, 5 for D, 9 for N, 10 for J, 6 for F, 10 for SU, 7 for SA in the word /‘aba/, and 7 cases were used for A, 6 for D, 11 for N, 8 for J, 6 for F, 9 for SU, 6 for SA in the word /‘ava/. All the 107 recordings considered for the articulatory and acoustic analysis were judged by 3 listeners and scored between 4 and 5 on a 0-5 adequacy scale.

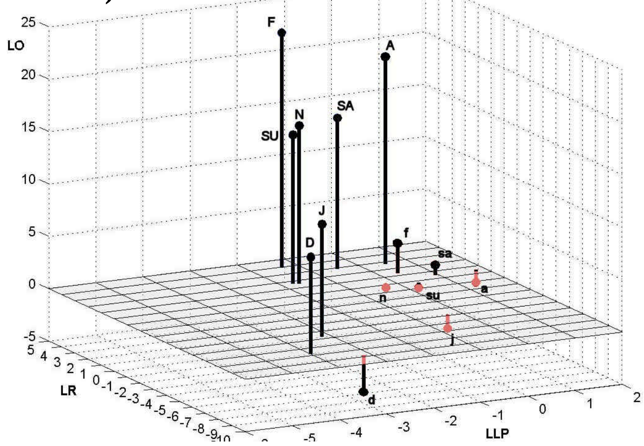
The speech signal, recorded synchronously with the lip movements, has been manually segmented and analyzed by means of a voice analysis software (PRAAT, [17]). Intensity, duration, spectrograms, formants, pitch synchronous F0, have been computed for /’aba/ and /’ava/, both for the whole word and for phonetic segments.

In this paper we present, for all the emotions introduced in § 2, the acoustic and articulatory parameters and the mean values characterizing the targets of the vowel /’a/ and the consonants /b/ and /v/, extracted with respect to the mid point of the corresponding acoustic signal. For what concerns the acoustic parameters, the F0 mean pitch and the spectral parameters F1, F2, F3 mean values were calculated in the mid point of the stressed vowels.

3. Articulatory data analysis

In Fig. 3a and 3b, the mean values of the LO, LR, and LLP in the stressed vowels (upper case labels) and consonants (lower case labels) of /’aba/ and /’ava/ are plotted in a 3D space. LLP is preferred over ULP since LL is the active articulator in the realization of the labiodental voiced fricative /v/ (see [1], [2]).

a) /’aba/



b) /’ava/

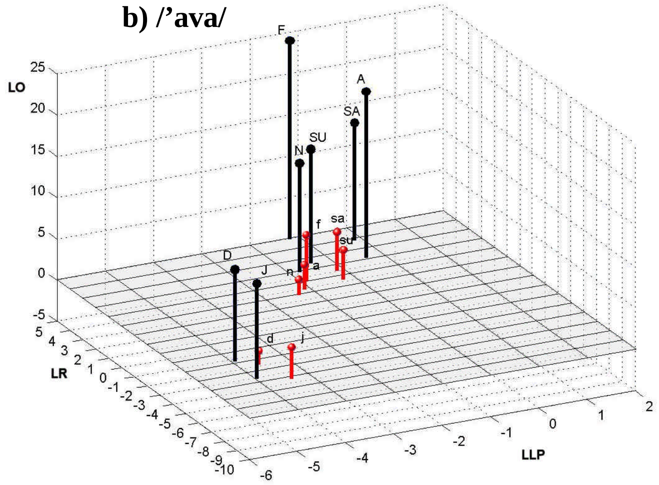


Figure 3: *Mean values of LO, LR, and LLP mid points defining the /’a/, /b/, and /v/ targets for /’aba/ and /’ava/. Upper case labels indicate vowel /’a/, lower case labels indicate consonants /b/ and /v/.*

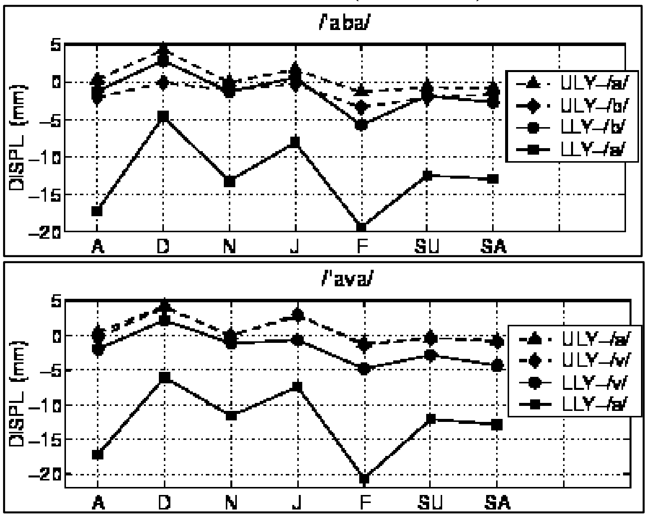


Figure 4*: Mean values of the UL and LL mid point vertical displacements for the /’a/, /b/, and /v/ in /’aba/ and /’ava/.*

The different contribution of UL and LL to the realization of LO is presented in Fig. 4a and 4b.

Fig. 5 shows the data referring to the horizontal (solid line ASYMX) and vertical asymmetries (broken line ASYMY) of the displacement of the corners of the lips.

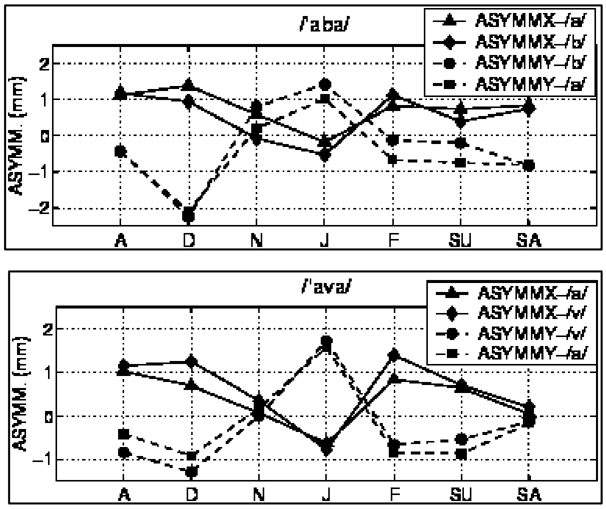


Figure 5*:Mean values of horizontal and vertical asymmetries of the lip corners for the /’a/, /b/, and /v/ in /’aba/ and /’ava/.*

A 2-factor ANOVA was performed to assess the effects of emotion (A, J, F, SA, D, SU, and also N) and stimulus (/’aba/, /’ava/) on each dependent articulatory variable for the vowel /’a/ and the consonants /b/ and /v/.

F-ratio associated to the main effects and interaction are reported in Table 1. For all parameters, the proportion of variance accounted by emotions was always significant.

For the vowel /’a/ emotion was the only significant effect, whereas for the two consonants the stimulus was important too.

Table 1: Results of the 2-factor ANOVA (emotion, stimulus, interaction) on articulatory parameters for vowel /’a/ and consonant /b/-/v/, F-ratio and p-value associated.

Param.	/’a/			/b/-/v/		
	Emot.	Stim.	Inter.	Emot.	Stim.	Inter.
LO	32.681*	0.179	0.826	8.002*	103.890*	1.012
LR	33.484*	0.940	0.673	63.725*	0.624	4.491*
UL	41.203*	2.641	0.862	19.488*	81.444*	2.860
LL	41.259*	0.097	0.888	20.926*	2.727	0.382
ULP	49.042*	0.106	1.061	63.146*	4.151	8.701*
LLP	23.424*	3.354	1.959	19.018*	133.397*	5.027*
ASYMX	8.319*	6.450	0.751	9.462*	0.107	0.621
ASYMY	15.151*	1.632	0.924	14.875*	0.001	1.583

* p<0.01

In order to understand the 2-factor ANOVA results the mean values of each parameter for emotion on vowel and consonants are examined separately.

3.1 Lip Opening

As regards the vowels /’a/ (see Fig. 3), with respect to neutral (N:15 mm), LO distinguishes significantly 3 groups of emotions: 1) the higest values are reached for fear (F: 22 mm) and anger (A: 19 mm); 2) sadness (SA) and surprise (SU) measure 14 mm and 13.9 mm respectively; 3) disgust and joy present the lowest LO displacement (D: 9 mm and J:10 mm).
From the 2-factor ANOVA, we can establish that the consonants distinguish significantly for stimuli apart from emotions. Concerning /b/, neutral presents a characteristic LO negative value (n: -0.4 mm) determined by the lip compression which is related to the articulatory closure phase. The highest negative value is reached by disgust (d: -3.5 mm), followed by joy (j: -1.3 mm) and anger (a: -1.2 mm), while fear and sadness stem out for they exceed the resting position by respectively 3 mm and 1 mm. As regards /v/, LO presents a limited range of positive values, determined by the labiodental constriction, which is the relevant phonological feature. In fact, LO varies from the lowest value for neutral (n: 2 mm) to the highest value for fear (f: 5.5 mm), and middle values for sadness (sa: 4 mm) and joy (j: 3 mm).

3.2 Lip Rounding

Concerning the two vowels /’a/ in Fig. 3, with respect to neutral value (N: 1.7 mm), LR reaches the highest positive values for fear (F: 4.4 mm), sadness (SA: 3.4 mm), anger (A: 2.4 mm) and surprise (SU: 2.1 mm), whereas the lowest negative values, corresponding to the lip spreading, are found for joy (J: -6 mm) and disgust (D: -6.2 mm).

As we can see from the 2-factor ANOVA, the consonants are distinguished significantly by stimuli, apart from emotions. For the consonant /b/, anger (a: 0.47 mm) and sadness (sa: 0.45 mm) determine higher rounding values than neutral (n: -0.3 mm), while joy presents intermediate values (j: -4.8 mm) and disgust is characterized by the maximum lip spreading (d: -8.8 mm).
For /v/, with respect to neutral production (n: 0.08 mm), anger (a: 0.3 mm), surprise (su: 0.5 mm), fear (f: 1 mm) and sadness (sa: 1.4 mm) present positive lip rounding values, while disgust and joy are characterized by the maximum lip retraction values (d: -5.4 mm and j: -7.3 mm).

3.3 Lower Lip Protrusion

As concerns /’a/ (see Fig.3), disgust and joy present LLP lowest values corresponding to lip retraction (D: -4.5 mm, J: -4.1 mm), whereas the other emotions are characterized by mean values between -1.7 mm for neutral and 0.3 mm for anger.
The consonants are significantly distinguished not only by emotions, but by stimuli too. For /b/, with respect to neutral mean value (n: 0.4 mm), anger (a: 1.8 mm) and sadness (sa: 0.9 mm) are characterized by the highest protrusion, and disgust (d: -3.2 mm) by the highest retraction. For /v/, anger, fear, sadness and surprise present negative values of LLP between -2.2 mm (a) and -1.2 mm (sa), while neutral is characterized by an intermediate value (n: -2.3 mm) and disgust and joy present the highest lip retraction values (d: -4.6 mm and j: -4.4 mm).

3.4 Upper Lip Displacement

UL vertical displacement values in the two vowels /’a/ change significantly according to the emotions (see Fig. 4). With respect to the neutral (N: -0.04 mm), anger, joy and disgust are characterized by positive values (A: 0.3 mm, J: 2.2 mm, D: 4.2 mm), corresponding to UL raising. In the other emotions, UL presents values between -1.2 mm for fear and 0.6 mm for surprise.
The mean values for consonants vary significantly not only by emotions, but by stimuli too. For /b/, UL has always negative values, between -0.1mm for disgust (d), -1.1 mm for neutral (n) and -3.4 mm for fear (f), corresponding to UL lowering.
In /v/ consonant, UL is not involved in labiodental constriction while it is subjected to the coarticulatory effects of the contextual vowel. Therefore disgust (d: 4 mm) and joy (j: 3 mm) are characterized by positive displacements which correspond to UL raising movements. The other emotions present negative values for anger and sadness between -0.3 mm (a, su) and -1.2 mm (sa, f), with the neutral characterized by an intermediate value of -0.06 mm.

3.5 Lower Lip Displacement

In LL displacements the vowels and the consonants vary significantly according to the emotions (see Fig. 4).
It is possible to define 3 groups for the stressed vowels: the first includes the emotions characterized by the widest displacement: fear (F: -20 mm) and anger (A: -17.2 mm; negative values correspond to the lip opening). The second group contains neutral (N: -12.4 mm), surprise (SU: -12.3 mm) and sadness (SA: -12.9 mm). The third group consists of joy and disgust (J: -8.3 mm and D: -5.2 mm), which present the lowest displacement values.
LL displacement mean values in /b/ and /v/ are similar and both narrower than the vowels as LL realizes the lip closure and labiodental constriction. In /b/ consonant, we found in fear (f: -5.8 mm) the greater displacement value, comparing to neutral (n: -1.3 mm), followed by sadness (sa: -2.8 mm). Positive values, corresponding to a LL raising, characterize disgust (d: 2.8 mm) and joy (j: 0.6 mm).
In /v/ consonant, fear (f: -4.8 mm) and sadness (sa: -4.3 mm) are characterized by the largest displacement mean values. Positive values are realized only for disgust (d: 2.2 mm).

3.6 Horizontal and Vertical Asymmetries

The 2-factor ANOVA shows for the stressed vowels and the consonants that emotions were the only significant effect (see Fig. 5). Vowels and consonants in all emotions (except joy) present a right horizontal asymmetry (ASYMX i.e. the right corner is displaced on the right side more than the left corner). Moreover we found a vertical left asymmetry (ASYMY i.e. the right corner is lower than the left one), except for joy with a left horizontal asymmetry and neutral where there is not asymmetry. Summarizing, joy is characterized, with respect to the other emotions, by a right vertical asymmetry and a left horizontal asymmetry.

4. Acoustic data analysis

In Fig. 6 a) and b), we can see the 3D plots of the mean values of F1, F2, F3 spectral characteristics, extracted in the stressed vowels mid points. Table 2 presents the mean values of F0 mid points for /'a/, /b/ and /v/. A 2-factor ANOVA was performed to assess the effects of emotions and stimuli on the chosen acoustic parameters. F-ratio associated to the main effects of emotions, stimuli and interaction are reported in Table 3. For vowels and consonants and for all parameters, except F1, the proportion of variance accounted by emotions was significant. As previously shown in Figure 6, the F1 values for /'a/ of /'aba/ vary between 794 Hz of surprise and 895 Hz of joy, while F1 values for /'a/ of /'ava/ vary between 780 Hz of surprise and 910 Hz of fear. Due to this limited range of variation, F1 was not distinguished significantly by emotions. On the contrary F2 values are significantly different by emotions. With respect to neutral (N: 1347 Hz), sadness presents the lowest value of F2 (SA: 1258 Hz), while we find the highest values for joy (J: 1522 Hz) and fear (F: 1786 Hz). As concerns F3, neutral is characterized by a mean value of 2585 Hz, while anger and disgust by values ranging between 2490 Hz and 2600 Hz. Joy, surprise and sadness present values ranging between 2700 Hz and 2800 Hz. Fear is characterized by the highest F3 values for /'a/ of /'aba/ (3163 Hz) while sadness for /'a/ of /'ava/ (3086 Hz).

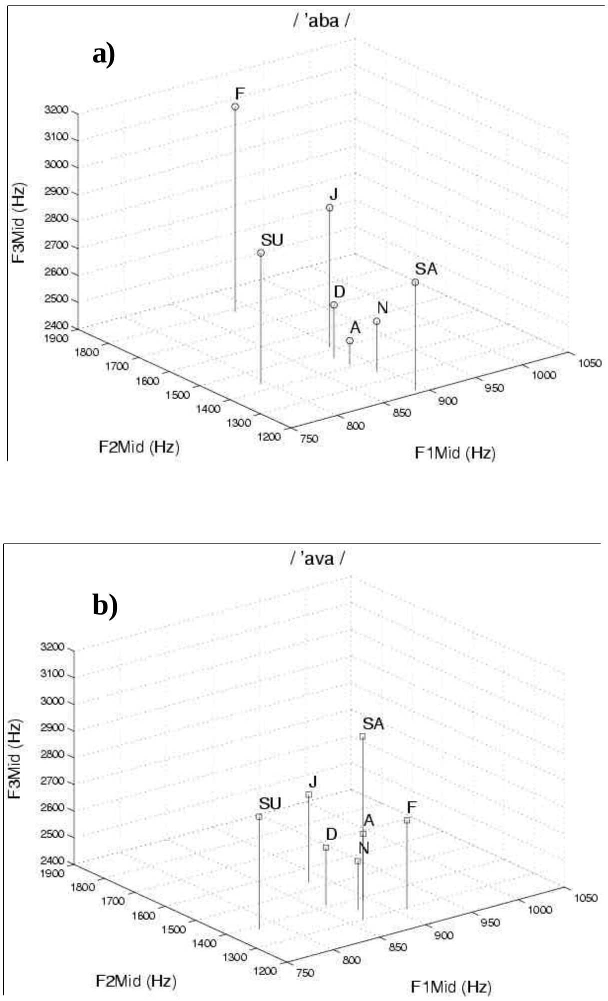


***Figure 6**: 3D representation of spectral characteristics F1, F2, F3 of the vocalic targets for all emotions and neutral production.*

Table 2: Mean values expressed in Hz of the F0 mid points for stressed vowels and consonants in /'aba/ and /'ava/.

	A	D	N	J	F	SU	SA
/'a/ in /'aba/	185.6	145.6	124.6	287.9	291.4	262.5	197.9
/'a/ in /'ava/	193.5	143.3	125.1	287.6	315.6	264.7	207.2
/b/	146.2	105.2	119.2	154.8	245.4	172.6	168.8
/v/	153.5	144.1	118.9	162.5	251.2	151.2	156.5

Table 3: Results for the 2-factor ANOVA (emotion, stimulus, interaction) on acoustic parameters for vowel /'a/ and consonant /b/-/v/, F-ratio and p-value associated. * $p<0.01$

Param.	/'a/			/b/-/v/		
	Emot.	Stim.	Inter.	Emot.	Stim.	Inter.
F0	21.887*	0.003	1.868	54.530*	0.525	1.229
F1	2.251	0.063	0.386			
F2	3.846*	4.510	4.078*			
F3	4.053*	1.322	1.966			

5. Correlations between Acoustic and Articulatory Data

In order to define the relationship between the acoustic and the articulatory parameters the correlations between the two series of data presented in §3 and §4 are calculated. In this paper we describe only the strong correlations, positive and negative, characterized by a score between 0.750 and 1.000. As concerns the neutral: LO, LR, ULP are negatively correlated to F0; LO and ULP are positively correlated to F1; LR is negatively correlated to F1[3]. For anger there are several strong correlations: LR and ULP are both strongly negatively correlated to F3; LO is strongly negatively correlated to F0; LR is strongly negatively correlated to F3 and ASYMX is strongly positively correlated to F0. As regards joy we found only one negatively strong correlation between LO and F2. In sadness ULP is negatively correlated to F1; disgust has only one negative correlation between F1 and ASYMX while in surprise ASYMY is strongly positively correlated to F1.

6. Conclusions

Although the data presented in this paper are only a first contribution to the issue of visual and acoustic segmental characteristic of emotive speech, several results are interesting. The modifications induced by emotions are pointed out for all labial kinematic and acoustic chosen parameters. We verified that some labial parameters vary significantly not only for emotions but for stimuli too: LO, UL and LLP values are distinguished by the consonants. On the basis of these quantifications it is possible to characterize visually and acoustically the emotions: i.e., disgust and joy are both characterized by LO and LL lowest displacement values and by the highest spreading and retraction values. However these two emotions are distinguished by means of the horizontal and vertical asymmetry values and also, as regards acoustic parameters, by means of F0, F2 and F3 values (always highest in joy, near to neutral values in disgust). Fear and anger are well characterized by the highest values of LO, LL, and LLP, but are distinguished by means of F0, F2 and F3 values (always higher for fear). Of course these first results will be integrated not only by the analysis of other speakers but also checked out by bimodal analysis-by-synthesis experiments.

[3] These correlations for the neutral are in accordance with previous results for the Italian language [19] and for a large number of languages [20].

7. References

[1] Magno Caldognetto E., Zmarich C., Cosi P. and Ferrero F. E., "Italian Consonantal Visemes: Relationship between Spatial/Temporal Articulatory Characteristics and Co-produced Acoustic Signal", *Proc. of AVSP'97*, C. Benoit and R. Campbell (Eds.), Rhodes (Greece), 1997, 5-8.

[2] Magno Caldognetto E., Zmarich C., Cosi P., "Statistical Definition of Visual Information for Italian Vowels and Consonants", in *Proc. of AVSP '98*, D. Burnham, J. Robert-Ribes and E. Vatikiotis-Bateson (Eds.), Terrigal (Aus), 1998, 135-140.

[3] Ekman P. and Friesen W., *Facial Action Coding System*, Consulting Psychologist Press Inc., Palo Alto (USA).

[4] Scherer K.R., "Vocal Affect Expression: A Review and a Model for Future Research", *Psych. Bull.*, 99, 1986, 143-165.

[5] Banse R. and Scherer K.R., "Acoustic Profiles in Vocal Emotion Expression", *Journal of Personality and Social Psychology* 170, 1996, 614-636.

[6] Ekman P., Friesen W. and Hager J.C., *New Version of the Facial Action Coding System*, A Human Face Publ., Salt Lake City (Utah), 2002 (eBook, Cd).

[7] Scherer K.R., "Vocal Communication of Emotion: A Review of Research Paradigm", *Speech Comm.*, 40, 2003, 227-256.

[8] Douglas_Cowie E., Campbell N. (Eds.), *Special Issue on Speech and Emotion*, in *Speech Comm.*, 40 (1-2), 2003, 1-258.

[9] Hess U., Kappas H., Scherer K.R., "Multichannel Communication of Emotion: Synthetic Signal Production", in K.R. Scherer (Ed.), *Facets of Emotion: Recent Research*, Lawrence Erlbaum Ass. Publ., Hillsdale (N.J.), 1988, 161-182.

[10] Massaro D. W., Egan P. B., "Perceiving Affect from the Voice and the Face", *Psych. Bull.*, 3, 1998, 1021-1032.

[11] de Gelder B., Vroomen J., Bertelson P., "Cross-modal Bias of Voice Tone on Facial Expression: Upper versus Lower Halves of a Face", *Proc. ICSLP '98*, Sydney, 1998, 93-96.

[12] de Gelder B., Vroomen J., "The Perception of Emotions by Ear and Eye", *Cognition and Emotion*, 14, 2000, 289-311.

[13] Schroeder M., Aubergé V., Cathiard M., "Can we Hear Smile?", *Proc. ICSLP '98*, Sydney, 1998, 559-662.

[14] Aubergé V. and Cathiard M., "Can we Hear the Prosody of Smile?", *Speech Communication*, 40, 2003, 87-98.

[15] Camras L.A., Holland E.A. and Patterson M.J., "Facial Expression", in M. Lewis and J. Havilland (Eds.), *Handbook of Emotions*, The Guilford Press, New York, 1993, 199-208.

[16] Palmer A., Richard D. and Strobeck C., "Fluctuating Asymmetry: Measurements, Analysis and Patterns", *Annual Review of Ecology and Systematics*, 17, 1986, 391-421.

[17] Boersma P., "PRAAT, a system for doing phonetics by computer", *Glot International* 5 (9/10), 1996, 341-345 (PRAAT web site: http://www.fon.hum.uva.nl/praat/).

[18] Drioli C., Tisato G., Cosi P. and Tesser F., "Emotions and Voice Quality Rendering Experiments", submitted to *Voice Quality: Funtions, Analysis and Synthesis*, Aug. 2003 Geneva.

[19] Ferrero F.E., Magno Caldognetto E., Vagges K. and Lavagnoli C., "Some Acoustic Characteristics of the Italian Vowels", *J. of Italian Linguistics*, 3, 1978, 87-95.

[20] Whalen D.H. and Levitt A. G., "The Universality of Intrinsic F0 of Vowels, *J. of Phonetics*, 23, 1995, 349-366.