

ABOUT THE RELATIONSHIP BETWEEN EYEBROW MOVEMENTS AND F₀ VARIATIONS

Christian Cavé¹, Isabelle Guaitella¹, Roxane Bertrand¹, Serge Santi², Françoise Harlay³, and Robert Espesser¹

1. Laboratoire "Parole et Langage", Université de Provence, Aix-en-Provence, France

e-mail: cave@romarin.univ-aix.fr

2. Laboratoire de Phonétique, Université de Franche-Comté, Besançon, France

3. Laboratoire de Neurobiologie Fonctionnelle, Université de Provence, Marseille, France

ABSTRACT

Speech production is always accompanied by facial and gestural activity. The present study is part of a broader research project on how head movements and facial expressions are related to voice variations in different speech situations. Ten normal subjects were recorded while reading aloud, answering yes/no questions, and dialoguing with an interviewer. Rapid rising-falling eyebrow movements produced by the subjects as they spoke were associated with F₀ rises in only 71% of the cases. This suggests that eyebrow movements and fundamental frequency changes are not automatically linked (i.e., they are not the result of muscular synergy), but are more a consequence of linguistic and communicational choices. Note also that 38% of the eyebrow movements were produced while the subject was not speaking. Thus, eyebrow movements may also serve as back-channel signals or play a role in turn-taking during conversation.

1. INTRODUCTION

Based on a multimodal approach to communication, this exploratory study is part of a broader research project on how head movements and facial expressions are related to voice variations during speech. It attempts to experimentally validate our hypotheses on the links between rapid eyebrow movements and changes in fundamental frequency during speech production in various situations. We noted in a previous study that rapid rising-falling eyebrow movement curves were associated with F₀ curves with the same shape (Cavé, Guaitella, & Santi, 1993). In the present study, raw gestural data was collected using an automatic movement analysis system. Once converted into curves, the gestural data was compared with the fundamental frequency curves output from an automatic signal processing system.

1.1 Trimodal Communication

The theoretical framework of the present study is based on a trimodal conception of communication wherein the verbal, the vocal, and the gestural are the three modalities of interpersonal expression (see Bally, 1925; Mehrabian, 1972; Cosnier, 1988). One of the perspectives in the study of gestures and their relationship to intonation is that the vocal system and the gestural system are more closely linked to each other than they are to the verbal system (Guaitella, 1995). Research on the links between these means of expression is scarce, however, and most available studies deal solely with one modality.

1.2 Gestures

Darwin (1874) considered gestures that express emotions to be the outcome of the interaction between physiological and functional constraints. More recently, Cosnier (1991) showed that eyebrow raising in particular plays a role in question asking. Ekman and Friesen (1975) investigated emotional factors by studying facial cues, while Birdwhistell (1970) and Condon (1976) showed how certain gestures play an accentuating role, one such gesture being rapid eyebrow movements. For these authors, rhythmic marking by means of gestures is associated with rhythmic marking at the syllabic level by means of prosody. As such, gestures are thought to perform a genuine function in both expression and meaning. More recently, Pentland and Darell (1994) mentioned that eyebrow raising might be associated with a specific intonation structure.

1.3 Voice-Gesture Links

A few studies have attempted to provide evidence of a connection between the vocal system and the gestural system that might attest to coordinated planning at the cognitive level. In particular, they have found a relationship between the rising and falling of the fundamental frequency, and the direction of the accompanying gestures (Morgan, 1953; Bolinger, 1985, Guaitella, 1991). Moreover, Fonagy's (1983) theory based on the work by Paget (1930) suggests that intonation is rooted in laryngeal kinesics, making intonation the acoustic reflection of a glottal movement.

For Bolinger (in press), the study of gestures and facial expressions should be conducted in relation to the study of intonation. Signs which associate intonation and gestures fulfill both a semiotic function and a rhythmic function (Guaitella, 1991, 1995). Gestures perform a dynamogenic function during speech encoding and a rhythmic function during vocal production. At the semiotic level, the relationship between gestural activity and vocal activity cannot be described as the mere sum of elementary meanings, but rather as the management of a global semiotic modality wherein each element loses its individual meaning. Thus, it is the combination of various signals emitted in parallel that determines the overall meaning of the message.

2. EXPERIMENTAL METHOD

Ten subjects volunteered to participate in the experiment without pay. They were recorded for one hour in each of three different speech production situations: (1) reading a short text aloud, (2)

answering ten yes/no questions, and (3) dialoguing with an interviewer about their work.

2.1 Data Acquisition

For each subject, three recordings were made in each speech situation. (1) The speech signal was recorded using a DAT recorder and a Sennheiser microphone (type ME 40). (2) The subject was filmed on videotape throughout the entire experimental session. (3) The head and face movements were recorded using an automatic movement analysis system.

Kinematic data acquisition and processing. Head and eyebrow movements were recorded using a movement analysis system (called Elite) with an automatic 100 Hz optical TV-image processor. The system reconstructs the three-dimensional trajectories of small infrared retroreflective markers (Ferrigno and Pedotti, 1985). The movements of the five points chosen for the present study were captured by means of five hemispherical markers (5 mm in diameter) glued onto the subject's skin: one marker on the forehead at the upper extremity of the frontal suture for recording head movement, one marker on each inside edge of the left and right supra-orbital arches for scowl detection, and one marker on each outside edge of the left and right supra-orbital arches for analyzing eyebrow raising and lowering. For the last two markers, a lateral rather than middle position was chosen to avoid confusion between the inside and outside markers during head rotation. An external reference marker was placed on a vertical rod behind the subject.

Two video cameras equipped with wide-angle lenses were placed on either side and to the front of the subject, at a distance of 1.5 m from his or her frontal plane.

The Elite system detected the markers by means of a real-time shape recognition procedure. It then calculated the x-y coordinates of the centroid, and reconstructed the trajectories in three-dimensional space. The powerful algorithms used in the centroid calculations and stereometric procedure attained a measurement accuracy of about 1/3000, i.e. 0.5 mm for the linear movements in our experimental conditions. After digital filtering based on preliminary automatic selection of the bandwidth adjusted to each signal (D'Amico and Ferrigno, 1990), the first and second derivatives of the linear displacements of the markers were computed.

Because the eyebrow movements were masked by the large head movements, they were isolated by means of a procedure that translated the (x, y, z) coordinates of the eyebrow markers, using the forehead marker as the origin (Berthoz and Pozzo, 1988). This made it possible to superimpose all frames without altering parameter estimates. The residual movements in the frontal plane could thus be attributed to the eyebrows. This method is valid since the forehead marker placed on the median line and upper part of the face was not affected by the eyebrow movements, and therefore necessarily expressed movement of the head.

2.2 Data analysis

The kinematic data output by the Elite system (x, y, and z coordinates of the eyebrow movements) and the speech signal were transferred to a Unix-operated SUN station. An automatic procedure was used to display the spectrogram of the speech signal, the corresponding Fo curve, and the eyebrow movement curve (see figure 1).

2.3 Selection of data for analysis

For this first stage of our data analysis, we examined the corpus of the interview dialogue of three subjects (C, E, and F) drawn at random from among the ten. The present study is limited to the kind of eyebrow movements already examined in previous research (Cavé, Guaitella, & Santi, 1993), i.e. rapid rising-falling movements. Inverted trajectories were eliminated (cases where the two eyebrows moved in opposite directions), leaving those instances where either both eyebrows moved in the same direction, or one moved and the other did not. The cutoff point for considering a movement was set at 3 mm or more for at least one eyebrow.

For each eyebrow movement selected, two measures were taken: the duration (in ms) of the rise (for joint eyebrow movement, the duration of the simultaneous rise was used) and the magnitude (in mm) of the movement of each individual eyebrow. The corresponding speech signal curves were then automatically isolated. Two cases were encountered: the speaker was either silent ("pauses") or he/she was producing speech. In the latter case, the fundamental frequency curve was classified into one of five categories: rising-falling (RF), falling-rising (FR), plateau (P), rising (R), and falling (F). Cases where the speech of the two individuals overlapped were discarded because the fundamental frequency could not be analyzed.

3. RESULTS

As a whole, the speakers exhibited a large degree of eyebrow movement variability, particularly in movement magnitude where E differed from C and F. The number of movements selected was 30 for C, 36 for E, and 12 for F, for corpora of comparable lengths (approximately 5 minutes). Note also that most of the cases eliminated due to inverted trajectories pertained to subject F. Substantial inter-subject differences were also found for the movement magnitude of the left and right eyebrows taken separately. In contrast, the subjects did not differ significantly in mean movement duration, which was 376 ms. for the three subjects pooled.

Rise duration was not correlated with movement magnitude. Thus, large movements can take place in a very short time.

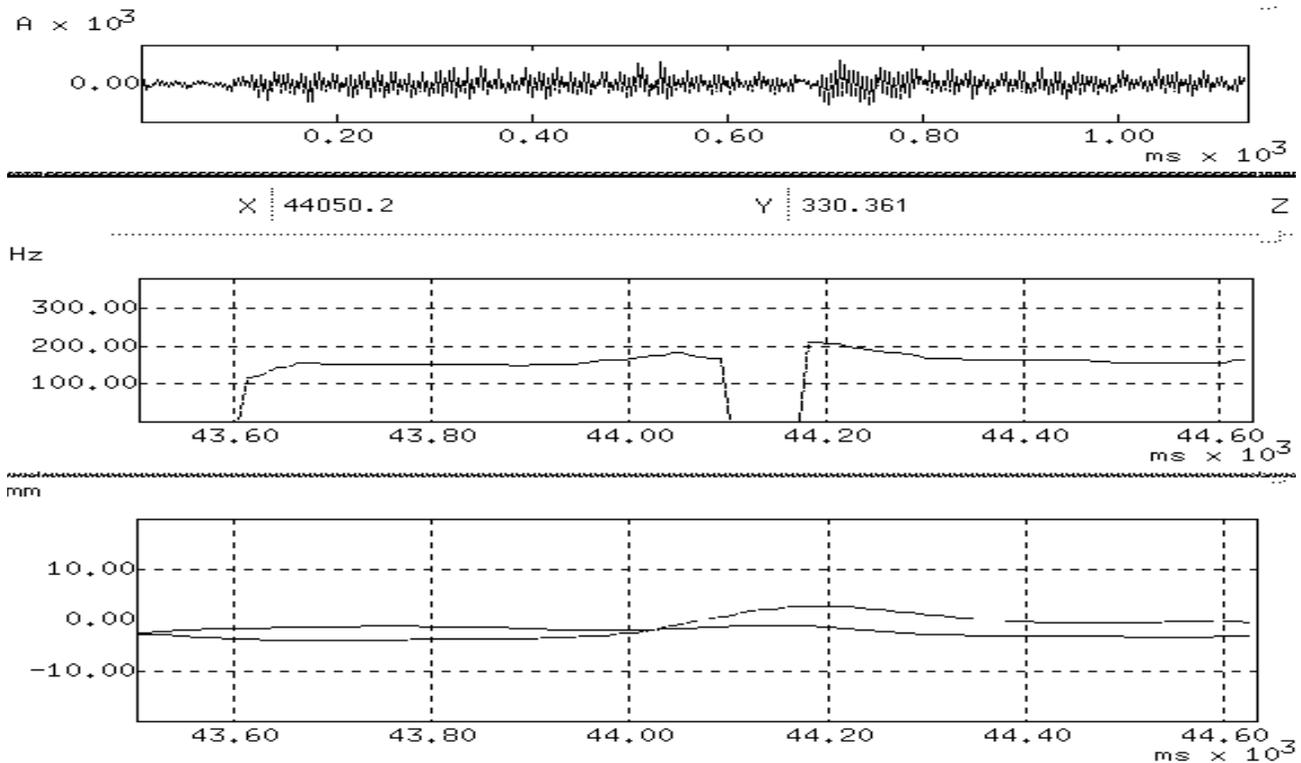


Figure 1. An example of curves displayed during data analysis. Upper: spectrogram of speech signal. Middle: F₀ curve. Lower: eyebrow movement curve.

Eyebrow movements during pauses represented 38% of the cases (see table 1). Further analysis is needed, however, to determine whether these gestures correspond essentially to back-channel signals or to movements that precede and assist the speaker in obtaining a speaking turn.

Speaker	Pauses	Speech	Total
C	15	15	30
E	6	30	36
F	9	3	12
Total	30	48	78

Table 1. Number of eyebrow movements during speech and pauses, by speaker.

Speaker	RF	FR	R	F	P	Total
C	10	1	2	1	1	15
E	12	0	6	2	10	30
F	0	2	1	0	0	3
Total	22	3	9	3	11	48

Table 2. Number of F₀ patterns of each type accompanying eyebrow movements, by speaker.

For each subject, table 2 gives the number of fundamental frequency patterns of each type accompanying eyebrow movements. The most frequent F₀ pattern was rising-falling.

Patterns containing a rise (RF, FR, and R) made up 71% of the instances of eyebrow movement during speaking.

The fundamental frequency pattern was not related to the duration of the eyebrow movements, nor to the magnitude of the movements of the right eyebrow, but was significantly linked to the movements of the left eyebrow. Large magnitude variations were found for the left eyebrow on the rising-falling patterns (coefficient of variation: 72.2). This did not occur for the right eyebrow, where the magnitude was more stable (coefficient of variation: 36.9).

There was no significant difference across subjects in fundamental frequency variation.

4. DISCUSSION

A great deal of behavioral variability across speakers was noted here for eyebrow movements but not for fundamental frequency. This may be linked to the situation under observation -- a dialogue about work with an unfamiliar interviewer -- which may have its own prosodic style. Eyebrow movements in this situation do not appear to be as stereotyped as in certain situations like news broadcasts on television.

The greater variation in movement magnitude for the left eyebrow may be due to functional differences between the brain hemispheres during communication. Differences between the left and right sides of the face during speech production are not

surprising. Lip opening, for example, is known to often be asymmetrical during speech production.

The stability of eyebrow movement duration (next to the highly variable movement magnitude and substantial overall behavioral differences across speakers) provides further evidence of the dichotomy between rapid and slow kinesic movements (Ekman, Friesen, 1975; Ekman, 1989).

The data obtained here are consistent with the results of previous research (Guaïtella, Cavé, & Santi, 1993) Indeed, Fo patterns with a rising component prevailed during rising-falling eyebrow movements. Eyebrow movements of this type were also observed during periods of speaker silence. What is new in the present study is the evidence of the presence of rising-falling eyebrow movements during flat or slightly rising pitch patterns (in particular, during hesitation phases). This suggests that eyebrow movements and fundamental frequency changes are not automatically linked (as they would be if they resulted from muscular synergy), but are more a consequence of linguistic and communicational choices. In the distribution of intonation patterns observed here, we again find the distinction between "rhythmic" or expressive movements which are associated with rises in the fundamental frequency, and "dynamogenic" movements, which assist in discourse planning and are associated with flatter Fo curves (Guaïtella, 1995).

The results of this preliminary study are encouraging, and provide incentive for future research on this subject.

5. REFERENCES

1. Bally, C., *Le langage et la vie*, Librairie Droz, Librairie Giard, 1925.
2. Berthoz, A., and Pozzo, A., "Intermittent head stabilization during postural and locomotor task in humans", In: B. Amblard, A. Berthoz, and F. Clarac (Eds.). *Posture and Gait: Development, adaptation and modulation*. Amsterdam, New York, Oxford: Excerpta Medica; 189-198, 1988.
3. Birdwhistell, R.L., *Kinesics and context*, University of Pennsylvania Press, 1970.
4. Bolinger, D.L., *Intonation and its parts*, London, Edward Arnold, 1985.
5. Bolinger, D.L. "Intonation in American English", in D. Hirst and A. Di Cristo (Eds) *Intonation systems: a survey of twenty languages*, Cambridge University Press (in press).
6. Cavé, C., Guaïtella, I., and Santi, S., "Fréquence fondamentale et mouvements rapides des sourcils: une étude pilote", *Travaux de l'Institut de Phonétique d'Aix*, vol. 15, 25-42, 1993.
7. Condon, W.S., "An analysis of behavioral organization", *Sign Language Studies*, 13, 285-318, 1976.
8. Cosnier, J., "Grands tours et petits tours". In Cosnier, Gelas, and Kerbrat-Orecchioni (Eds), *Echanges sur la conversation*, Editions du CNRS, Lyon, 154-84, 1988.
9. Cosnier, J., "Les gestes de la question", In Kerbrat-Orecchioni (Ed), *La question*, Presses Universitaires de Lyon, 163-71, 1991.
10. D'Amico, M., and Ferrigno, G., "Technique for the evaluation of derivatives from noisy biomechanical displacement data using a model-based bandwidth-selection procedure", *Med. and Biol. Eng. and Comp*; 28, 1990.
11. Darwin, C., *L'expression des émotions chez l'homme et les animaux*, French translation by S. Pozzi and R. Benoit, Reinwald and Cie Libraires-Editeurs, 1874.
12. Ekman, P., "La mesure de l'expression faciale", *Sciences et Vie*, 168, 24-31, 1989.
13. Ekman, P., and Friesen, W., *Unmasking the face. A guide to recognizing emotions from facial clues*, Prentice-Hall, 1975.
14. Ferrigno, G., and Pedotti, A., "Elite: A digital dedicated hardware system for movement analysis via real-time TV signal processing", *IEEE Transactions on Biomedical Engineering*, BME-32, II, 943-949, 1985.
15. Fonagy, I., *La vive voix*, Paris, Payot, 1983.
16. Guaïtella, I., *Rythme et parole: comparaison critique du rythme de la lecture oralisée et de la parole spontanée*, Doctoral dissertation, Aix-en-Provence, 1991.
17. Guaïtella, I., "Mélodie du geste, mimique vocale?" *Semiotica*, 103, 3/4, 253-276, 1995.
18. Guaïtella, I., Cavé, C., and Santi, S., "Relations entre geste et voix: Le cas des sourcils et de la fréquence fondamentale", In *Images et Langages: Multimodalité et modalisation cognitive. Actes du colloque interdisciplinaire du Comité National de la Recherche Scientifique*, Paris, 261-268, 1993.
19. Mehrabian, A., *Silent message*, Wadsworth, Belmont, 1972.
20. Morgan, B.Q., "Question melodies in American English", *American Speech*, 2, 181-191, 1953.
21. Paget, R., *Human speech*, New York: Harcourt-Brace, 1930.
22. Pentland, A.P., and Darell, T., "Visual perception of human bodies and faces for multi-modal interfaces", *Proceedings of the ICSLP 94*, Yokohama, S11-5.1, 543-546, 1994.