EYEBROW MOVEMENTS AND VOICE VARIATIONS IN DIALOGUE SITUATIONS: AN EXPERIMENTAL INVESTIGATION

Christian Cavé*, Isabelle Guaitella, and Serge Santi

Laboratoire Parole et Langage, UMR CNRS 6057, Université de Provence, Aix-en-Provence, France
*email: christian.cave@lpl.univ-aix.fr

ABSTRACT

This paper deals with the relationships between rapid eyebrow movements and vocal production during dialogue. A sophisticated movement and voice recording and analysis system was used to determine when eyebrow movements occurred relative to the phases of phonation, and how these movements were linked to accentuation-related intonative variations. The results confirmed the previously observed association between rapid eyebrow movements and F0 changes. Temporal relationships between eyebrow movements and the beginning and end of a speaker’s vocalization suggest that eyebrow movements act as a turn-taking cue.

1. INTRODUCTION

Taking a multimodal approach to communication, we experimentally investigated the links between rapid eyebrow movements and phonation during dialogue. In previous experiments, we examined the relationship between eyebrow movements and variations in the fundamental frequency of the voice (F0). We found event simultaneity and resemblance of form in the eyebrow-movement and fundamental-frequency curves [1, 2]. In the present study, special attention is paid to the temporal relationships between the beginning and end of phonation and eyebrow movements, in an attempt to determine the potential role of eyebrow movements as turn-taking and/or back-channel signals. For the data presented here, we used an experimental device that supplies objective gestural data via an automatic movement acquisition and analysis system (Elite). Once the eyebrow data is plotted on curves, it can be related to and compared with the curves obtained using sound signal processing software.

2. THEORETICAL BACKGROUND

Rapid eyebrow movements, also known as eyebrow flash, were described by Eibl-Eibesfeldt [3] as a part of ritual greetings. Scientists like Ekman and Friesen [4] studied facial cues to gain insight into the emotional aspects of communication, and Ekman in particular attempted to characterize rapid movements, of which eyebrow movements are a part. Such rapid movements are thought to be mainly linked to discourse production, where according to Birdwhistell [5] they serve as a rhythmic demarcation device. For many authors, the difference between rapid and slow movements is a categorical one, with rapid movements being seen as rhythmic and slow movements being viewed as semiotic. We have criticized this point of view [6, 7], and contend that both rapid and slow movements contribute not only to rhythm, but to semiotic production as well.

Rapid eyebrow movements have a semiotic component that is rooted -- probably by association -- in the emotional expressions of the face. However, their function is essentially discourse-based, insofar as their use in the expression of emotions has come to take on a linguistic function. For example, an eyebrow movement occurring on an unpredictable word can be interpreted as "Pay special attention to the word I'm pronouncing." For Bolinger [8], the eyebrows even more than other types of gestures are related to a privileged way to speech, and benefit from a status equivalent to vocality.

3. HYPOTHESES

Our first hypothesis was that eyebrow movements play a role in managing the conversational interaction, i.e., in turn-taking. It was assumed that speakers would use this device to signal their desire to say something and thus, to be given a speaking turn. This hypothesis was tested by measuring the interval between the eyebrow movements and the beginning and ending of the speaking turns. Eyebrow movements that fell closer to the onset of a speaking turn than to its offset would lend support to this hypothesis.

Our second hypothesis concerned the role of the eyebrows in rhythmic demarcation. During speech production, a speaker may make an eyebrow movement at the same time as he or she stresses a syllable (of an important word, for example) by means of an intonation variation (by using a fundamental frequency contour with an "accentuating" value, i.e., one likely to lead to the perception of a stressed unit). This hypothesis was tested in a previous analysis of video documents [1, 2], which revealed a strong link between rapid upward-downward eyebrow movements and the fundamental frequency. The F0 curve had the same rising-falling contour (and thus, an accentuating value).
as the eyebrow movement curve, and the two curves were synchronized. The questions raised by this study were whether this link is truly systematic, or even automatic, and whether it is rooted in the effects of muscular synergy and/or is the outcome of a communicational strategy on the part of the speaker.

4. METHOD

4.1. Corpus

Ten unpaid subjects volunteered to participate in the experiment. They were questioned about their work by means of informal interviews conducted in a highly conversational style. The corpus consisted of ten dialogues, each lasting approximately five minutes. For technical reasons, only seven of the ten dialogues were processed.

4.2. Data acquisition

For each subject, three types of recordings were obtained: (1) the speech signal, (2) the video image, and (3) head and face movement data. Subjects were filmed on videotape for the full duration of the experimental session. The subject's speech was recorded using a DAT recorder and a Sennheiser microphone (type ME 40). Head and eyebrow movements were recorded using a motion analysis system (Elite) enabling computer reconstruction of three-dimensional trajectories of small infrared retro-reflective markers [9]. The kinematic properties of five key points were captured by means of small hemispherical markers attached to the subject's skin. The marker for recording head movements was placed on the forehead at the upper extremity of the frontal suture. For scowl detection, a marker was placed at each of the inner ends of the left and right supra-orbital arches. Markers for analyzing eyebrow raising and lowering were placed at the outer ends of the left and right supra-orbital arches. For these last two markers, a lateral position was chosen to avoid confusion between the inner and outer 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 positioned in front of the subject parallel to his/her frontal plane at a distance of 1.5 m. The sight angles of the two cameras in the horizontal plane were -35˚ and +35˚, respectively.

The Elite system detected the markers in real-time by means of a hardware shape-recognition procedure. It then calculated their centroid x-y coordinates and generated a 3-D reconstruction of the marker trajectories. The powerful algorithms used for the centroid calculation and the stereometric procedure allowed us to attain a measurement accuracy of about 1/3000, i.e., 0.5 mm for the linear displacements in our experimental conditions. After digital filtering based on the preliminary automatic selection of the appropriate bandwidth for each signal [10], the first and second derivatives of the markers' linear displacements were computed.

4.3. 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 Sparc SUN station in a Unix environment. An automatic procedure was used to display the spectrogram of the speech signal, the corresponding F0 curve, and the eyebrow movement curve. In order to separate the real movements of the eyebrows from those resulting from the movements of the head (which also affect the eyebrow sensors), the head movement component was factored out by an automatic procedure. The entire output dataset was verified by viewing of the video documents.

4.4. Selection of eyebrow movements

We limited the present study to rapid upward-downward eyebrow movements. Only those movements where both eyebrows moved simultaneously were retained for analysis. The cutoff point for inclusion of a movement was set at a displacement of at least 3 mm by at least one eyebrow.

4.5. Interval coding

To relate the occurrence of eyebrow movements to the vocal production, various intervals were measured, as follows:

(a) If the eyebrows moved when the speaker was talking, two intervals were measured: the time between the onset of the vocal signal and the eyebrow movement, and the time between the eyebrow movement and the end of the vocal signal (see Figure 1). These cases -- where the eyebrow movement occurred during phonation -- were put in Category 1 (C1).

(b) If the movement took place while the speaker was silent (even if the interviewer was talking), the intervals measured were between the eyebrow movement and the end of the speaker's last speaking turn, and between the eyebrow movement and the onset of the speaker's next speaking turn (see Figure 2a). Whenever the eyebrow movement and the end of the vocal production were synchronized (difference of 30 ms or less), the interval between the movement and the beginning of the next vocalization was measured. Similarly, whenever the movement was synchronized with the beginning of the vocalization (difference of 30 ms or less), the time since the end of the preceding vocalization was measured (see Figures 2b and 2c). These cases -- where the eyebrow
movement took place during speaker silence or at a vocalization boundary -- were grouped into Category 2 (C2).

4.6. Fundamental frequency coding
For the speaker's vocal production, the fundamental frequency contours were classified into seven categories: rising-falling (RF), falling-rising (FR), plateau (P), rising (R), falling (F), rising-falling-rising (RFR), and falling-rising-falling (FRF). In the passages where the speaker's and the interviewer's speech overlapped, the fundamental frequency could not be analyzed, so they were discarded. An intonation contour was considered to have an accentuating value when it contained points where there was a change in direction, called target points. Simple contours R and F were checked at the beginning and end to determine the presence of target points within the observed sequence. Only flat and falling contours had no target points, and thus no accentuating value.

5. RESULTS

5.1. Eyebrows and phonation
For Category C1 (movement during phonation), the eyebrow-movement/end-of-phonation interval averaged 2100 ms and the beginning-of-phonation/eyebrow-movement interval averaged 1500 ms. For category C2, the average end-of-phonation/movement interval was 1200 ms and the average movement/beginning-of-phonation was 490 ms.

For the two categories pooled, the interval data as a whole showed that eyebrow movements occurred closer to the beginning of a vocal production than to the end. This effect was not very strong but statistically significant (p < .05).

5.2. Eyebrows, accentuation, and fundamental frequency
Only C1 data was analyzed here. This category contained 43.5% of the observed eyebrow movements (excluding the 16 cases of overlapping out of 96). RF intonation contours -- which are similar in shape to eyebrow movement curves -- were found for 67.5% of the movements. Contours RF, FR, RFR, FRF, and R, all of which are accentuating, accompanied 93.75% of the movements. This left a mere 6.25% for contours with a non-accentuating value (falling: 5%, plateau: 1.25%).

6. DISCUSSION
The observed proximity of phonation onset and eyebrow movement supports the idea that eyebrow movements act as signals of a new speaking turn. Concerning the relationship between eyebrow movements and fundamental frequency, the data obtained here refine the results of previous studies [1, 2]. Although 93.75% of the eyebrow movements were associated with accentuating intonation contours (mostly RFs), the presence of eyebrow movements during non-accentuating productions (in particular, during periods of hesitation) seems to argue for the idea that the link between eyebrow movements and fundamental frequency variations is not an automatic one. In other words, eyebrow movements are not solely dictated by F0 variations, but may be speaker-controlled, being the consequence of linguistic and communicational choices. However, hand-voice coordination during the production of accentuating markers is known to be highly prevalent. The muscular synergy hypothesis cannot be rejected, then, even if it does not apply in all cases. In the contour distribution obtained here, we can see a distinction between what might be called "rhythmic" movements, which are more or less expressive and associated with accentuating variations in the fundamental frequency, and movements we could call "dynamogenic", which contribute to discursive planning and are associated with flatter contours [7].

Thus, rapid raising/lowering of the eyebrows does indeed appear to act as an attention-getting device, whether during the speaker's silent periods ("Look, I would like to speak"), or during speech production ("Listen, what I'm saying is important"). Accordingly, we can contend that this gesture in fact derives from the eyebrow-raising movement, whose primary function was to manifest enhanced visual attention.

In the future, we plan to look at what goes on for the interviewer, from the vocal and visual standpoints, when the speaker's eyebrows move, in hopes of finding out whether these attention signals may also function in a more interactive way, and at least in certain cases, reflect either focusing on the interviewer's speech or between-speaker rhythmic harmonization, as proposed in Condon's [11] intersynchrony hypothesis.

7. REFERENCES


Figure 1. Diagram of temporal relationships between eyebrow movements and phonation: eyebrow movement during phonation.

![Figure 1](image1.png)

Figure 2. Diagram of temporal relationships between eyebrow movements and phonation: eyebrow movement outside of phonation or at a phonation boundary.

2a. Eyebrow movement during silence

![Diagram 2a](image2a.png)

2b. Eyebrow movement occurring exactly at the end of phonation

![Diagram 2b](image2b.png)

2c. Eyebrow movement occurring exactly at the beginning of phonation

![Diagram 2c](image2c.png)