Facial Gestures and Communication: what induces raising eyebrow movements in Map Task dialogues

María Luisa Flecha-García

marisa@ling.ed.ac.uk

ABSTRACT

There is some evidence suggesting that eyebrow movements may have linguistic functions in communication, but research in this area has been scarce. This study tested the hypothesis that eyebrow raises have communicative functions and therefore will tend to be distributed unequally across conversational move types. The data consisted of six dialogues of participants performing a cooperative task, namely the Map Task. There was an overall significant difference in the distribution of eyebrow raises across move types. A preliminary analysis on one of the dialogues suggested that their distribution did not depend simply on differences in the length of the various move types. Further research is in progress that may eventually provide useful information for more efficient engineering of multimodal communication systems.

1 INTRODUCTION

The study of facial movements has attracted the interest of researchers from many different fields. One aspect that has received special attention is the expression of emotion on the face. But the face also encodes more properly linguistic messages. The articulatory movements of speech production are concentrated on the lower part of the face and have been studied, for instance, in visual speech recognition and audiovisual speech perception. However, research on upper-face movements in connection with speech is still scarce.

There is some evidence for the linguistic functions of upper-face gestures. One area in which this has been investigated is the study of signed languages, where it is widely accepted that some of these nonmanual markers are used to fulfil prosodic functions (Wilbur 1994; Corina 1989). Facial displays have been found to mark the introduction of a topic, and of certain clauses or questions. Obviously, the use of these displays with linguistic functions in signed languages is much more systematised than in the speech of hearing people. But it is nevertheless important to note how upper face movements can substitute for the auditory suprasegmental cues that signers do not have. Evidence for linguistic functions of non-articulatory facial gestures has also been found in spoken language. Ekman (1979) pointed out how two of the brow actions described in the Facial Action Coding System (Ekman and Friesen, 1978) seem to have important linguistic roles, such as: giving emphasis to words and phrases, marking punctuation, and signalling questions. Chovil (1991/92) studied videos of spontaneous dialogues in order to examine how facial displays are used in conversation and she described both syntactic and semantic linguistic functions in facial displays. Of the syntactic displays the most common were eyebrow movements that served as emphasisers and question markers. Looking at eyebrow movements in particular, Cave et al. (1996) compared the curves of rapid rising-falling eyebrow movements with the fundamental frequency curves of the accompanying speech. They found that eyebrow movements and

fundamental frequency changes were not automatically linked, suggesting that they were not the result of muscular synergy, but rather a consequence of linguistic and communicational choices.

In short, there have been some studies suggesting the linguistic functions of eyebrow movements but there is not enough knowledge about the way in which these gestures are combined with speech. This lack of information becomes a limitation in the design of multimodal communication systems. This paper presents a preliminary study of the relationship between different types of "moves" in conversation (Carletta *et al.* 1997) and the distribution of eyebrow movements. The hypothesis was that eyebrow raises have communicative functions and therefore they will tend to be distributed unequally across conversational move types. The aim of the project is to provide some information about audiovisual language production that could be used as a guideline for more efficient engineering of multimodal communication systems.

2 METHODOLOGY

2.1 Materials

To obtain visual and auditory data from spontaneous conversation, four female native speakers of British English were videorecorded performing a collaborative task, namely the Map Task (Anderson *et al.* 1991). In the Map Task, two participants, the Instruction Giver and the Instruction Follower, sit opposite each other with slightly different versions of a simple map. The Giver's map has a route and a set of landmarks, whereas the Follower's has no route. Their task is to reproduce the route on the Follower's map. Since the maps are not quite identical and they cannot see each other's maps they have to negotiate in order to reach the Finish Point. To allow a comparison between their behaviour in dialogues and monologues, the participants in this study were recorded performing the task as described above, but also alone describing the route to the camera. In the present study only the dialogues were analysed. Nevertheless, the whole design is described below.

2.2 Design and Procedure

Participants were taken into a recording studio and were recorded in four different situations:

2.2.1 Rehearsal

Each participant was instructed to do a rehearsal in front of a camera, describing the route on the map so that someone could reproduce that route from the video of her instructions. She was told that this would not be recorded but was necessary for the camera and sound to be adjusted as well as for her to get used to the task. However, both sound and image were recorded, with a frontal view of her upper body.

2.2.2 Recording of second monologue

Participants were asked to describe the route as in the rehearsal, but this time for a real video session addressed to a potential viewer.

2.2.3 Dialogue

Here participants were recorded in pairs as in the original design of the Map Task, and they had to collaborate to reproduce the Giver's route on the Follower's map. There were a total of eight dialogues in which four different maps were used. Each participant served as a Giver for the same route to two different Followers and as Follower for two different routes.

2.2.4 Reading of list

Finally, each participant was recorded reading a list of the landmark names that appeared in the different maps.

2.3 Annotation

Two different types of annotation were performed and then combined into a single text file.

2.3.1 Conversational Game Moves

Move annotation, the lowest level of dialogue structure as described by Carletta *et al.* (1997), was performed for all dialogues using Xwaves and xlabel from Entropic. The type of move known as *ready* in Carletta et al.'s coding scheme was not coded in this study. Also, in order to get sufficient number of cases, move types were grouped into larger categories as follows:

Current study	*Carletta et al.(1997)*
Instruct	Instruct
Explain	Explain
Query	Query-yn
	Query-w
	Check
	Align
Reply	Reply-y
	Reply-n
	Reply-w
Other	Acknowledge
	Clarify
	Unclassifiable

2.3.2 Eyebrow movements

Rising eyebrow movements from three speakers in the role of Giver were coded for a total of six dialogues. The beginning and end points of these movements were annotated on the timeline of the dialogue, without differentiating minor movements from more prominent ones. This annotation was performed using scripts originally written to code drawing times from Followers in a previous Map Task study. It allowed writing times into a file, while watching the video, by clicking the mouse every time speakers raised their eyebrows. This tool was the best available for this project at the time, but it has obvious limitations for the study of gesture because it did not allow absolute synchronisation of video and audio data.

2.4 Sample size

The data analysed here consist of three speakers (in the role of Giver) and six dialogues from which a total of 736 moves were extracted: 314 *instruct*, 65 *explain*, 93 *query*, 121 *reply* and 143 *other*.

3 RESULTS

Figure 1 shows what proportion of the conversational moves of any type have eyebrow raises in them (with the total number of moves of that type at the base of each bar). Overall there was a significant difference among rates of eyebrow movements across different move types ($p < .0001$).

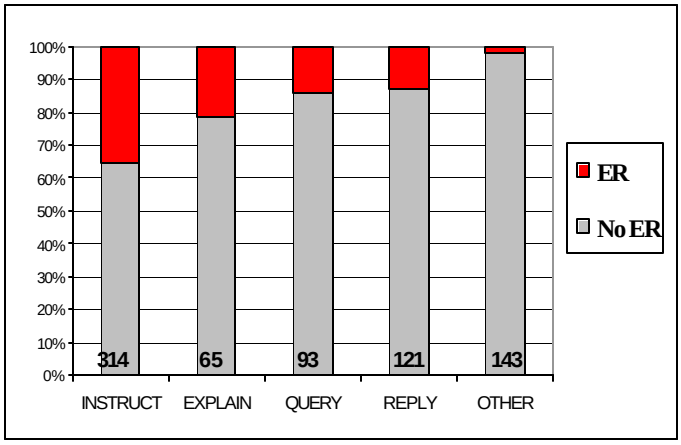

Figure 1. Rate of Eyebrow Raises (ER) across moves

The significant pairwise differences ($p < .005$) were as follows:
instruct > *reply*
instruct > *other*
explain > *other*

where *instruct* moves had significantly more eyebrow movements than *reply* and *other*, and *explain* moves had significantly more eyebrow movements than *other*.

The results above could indicate that eyebrow movements are distributed unequally across move types because they have particular communicative functions. However it could also be that eyebrow movements are randomly distributed and they simply tend to occur in move types that are generally longer. To explore this possibility, a small analysis was performed on one of the dialogues containing 94 moves (45 *instruct*, 16 *explain*, 9 *query*, 12 *reply* and 12 *other*).

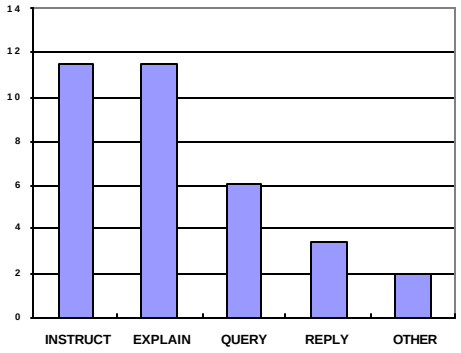

Figure 2. Average length (in number of words) of move types

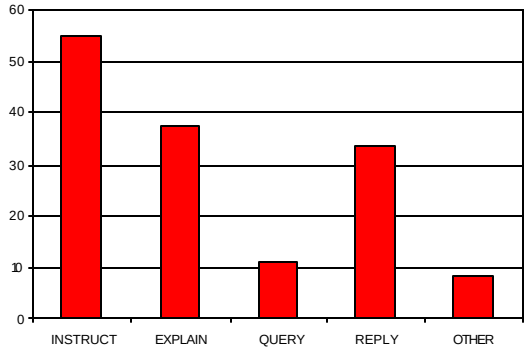

Figure 3. Rate of eyebrow raises across move types

Figure 2 shows the average length (in number of words) of the move types used in this analysis. And Figure 3 presents the percentage of eyebrow raises in each move type for the same dialogue. If the rate of eyebrow raises across move types depended on the length of the moves then we would expect both graphs to have a similar profile. However, the difference

in rate of eyebrow raises between *instruct* and *explain* in Figure 3 (53.3 and 37.5 respectively) does not appear in Figure 2, where both move types have practically the same average length (11.52 and 11.5, respectively). Similarly, in Figure 2 *query* is longer than *reply* (6.11 compared to 3.41), whereas in Figure 3 *query* has a smaller percentage of eyebrow movements than *reply* (11.11 compared to 33.33). These results suggest that, at least in the dialogue analysed here, the distribution of raised eyebrows across conversational moves was not related to the length of the moves.

4 DISCUSSION

This study tested the hypothesis that eyebrow raises have communicative functions and therefore they will tend to be distributed unequally across conversational move types. The results showed that there were significant differences in the percentage of raised eyebrows across different types of conversational moves. It was also found that, at least for one dialogue, the distribution of these eyebrow movements did not depend simply on differences in the length of the conversational moves. These results are preliminary and do not allow firm conclusions about the relation between eyebrow movements and dialogue structure. Nevertheless they clearly suggest that eyebrow raises are not randomly distributed and are associated with some types of conversational moves, such as *instruct* or *explain.*

One of the limitations in this study was the sample size (3 speakers and 6 dialogues). This is a common limitation in research on gesture in general, because the analysis of this type of data is very time consuming. Further analysis is currently in progress and includes the fourth speaker in the recordings as well as the monologues described in the design above. Another limitation was the tool employed to annotate the eyebrow movements on the timeline of the dialogues. A different tool (MacLaughlin *et al.* 2000) has been adopted for the analysis currently in progress that will provide more accurate synchronisation of the visual and auditory data. Future research will continue to investigate possible communicative functions of eyebrow movements, but it will also involve suprasegmental phenomena by looking at the distribution of eyebrow raises across pitch-accented words. Although this research is clearly exploratory, it may eventually provide some information about audiovisual language production that could be used as a guideline for more efficient engineering of multimodal communication systems.

ACKNOWLEDGEMENTS

Many thanks to my supervisors, Dr. Ellen G. Bard and Prof. Robert Ladd

References

Anderson, A. H., M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. S. Thompson & R. Weinert. 1991. The HCRC Map Task Corpus. *Language and Speech*, 34: 351-366.

Carletta, J., A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon & A. Anderson. 1997. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23:13-31.

Cave, C., I. Guaitella, R. Bertrand, S. Santi, F. Harlay & R. Espesser. 1996. About the relationship between eyebrow movements and Fo variations. *Proceedings of the ICSLP* (pp. 2175-2179). Philadelphia, PA, USA.

Chovil, N. 1991/92. Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction*, 25: 163-194.

Corina, D. P. 1989. Recognition of affective and noncanonical linguistic facial expressions in hearing and deaf subjects. *Brain and Cognition*, 9: 227-237.

Ekman, P. 1979. About brows: Emotional and conversational signals. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog, (eds.), *Human Ethology.* Cambridge: Cambridge University Press, pp. 169-248

Ekman, P. & W. V. Friesen, 1978. *The Facial Action Coding System (FACS): A technique for the measurement of facial action.* Palo Alto, CA: Consulting Psychologists Press.

MacLaughlin, D., C Neidle & D. Greenfield. 2000. SignStream (Version 2.0) [computer software and manual]. Retrieved from http://www.bu.edu/asllrp/SignStream

Wilbur, R. B. 1994. Eyeblinks and ASL phrase structure. *Sign Language Studies*, 84: 221-240.