# Head Movements, Facial Expressions and Feedback in Danish First Encounters Interactions: A Culture-Specific Analysis

Patrizia Paggio and Costanza Navarretta

University of Copenhagen, Centre for Language Technology (CST)
{paggio,costanza}@hum.ku.dk

**Abstract.** This study deals with non-verbal behaviour in a video-recorded and manually annotated corpus of first encounters in Danish. It presents an analysis of head movements and facial expressions in the data, in particular their use to express feedback, and it discusses the results in the light of aspects of Danish culture that seem to privilege rather unconventional and non-emotional behaviour. The data provided can form the basis of multi-cultural studies where parallels are drawn to similar interactions in other languages.

**Keywords:** head movements, facial expressions, gestural feedback, cultural differences in multimodal interaction.

## 1 Introduction

This study presents an analysis of how head movements and facial expressions are used in a video-recorded and manually annotated corpus of first encounter interactions in Danish with specific regard to the expression of feedback. It also discusses how aspects of the non-verbal behaviour observed in the corpus can be explained in terms of cultural specificity, and it proposes questions to be investigated from a multicultural perspective.

Many studies converge in showing that head movements correlate with different communicative functions. Nods are typically examples of backchannels, i.e. feedback signals given by the listener without trying to take the floor (Yngve, 1970; Duncan, 1972 and McClave, 2000), but they are also used in turn shifts (Hadar et al., 1984; 1985). In a subset of the Swedish GSLC corpus, it has been observed that 70% of all head movements are related to feedback, and that most of these are nods and up-down movements (Cerrato, 2007).

In our previous work (Jokinen et al. 2008), we studied facial expressions, head movements and hand gesturing in Danish and Estonian dialogues, and noticed significant interdependences between non-verbal expressions and communicative functions. Nods often indicate feedback, while head movements sideways or up-down together with gaze are related to turn-taking. Moreover, in Estonian dialogues there is a significant tendency to elicit feedback by looking at the partner, while acceptance is often accompanied by looking down. When clustering Danish facial data on the basis of the annotation features, we noticed that three clusters appeared in the data: one for eliciting

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2011, LNCS 6766, pp. 583–590, 2011.

feedback, one for acceptance, and one for contact/perception, thus indirectly providing evidence that the three functions are important in multimodal feedback analysis.

In this study, we look at a corpus of first encounter interactions in Danish created in the Nordic NOMCO project (Paggio et al. 2010). Our goal is assessing how head movements and facial expressions are used in the interactions, especially concerning the expression of feedback, and we discuss how these findings can be related to certain aspects of Danish culture.

The paper is structured as follows: in section (2) we review a number of previous studies of multimodal behaviour in different languages, and provide a characterisation of Danish with respect to a number of other cultures; in section (3) we present our expectations concerning the multimodal behaviour and we describe the data collection; in section (4) we describe the results; finally, in section (5) we conclude.

## 2 Multimodal Feedback in Different Cultures

Cultural differences exist in the way multimodal behaviour is used to express feedback, and a number of studies have already pinpointed specific characteristics of different languages. Maynard (1987), for example, studies head nods in dialogues between Japanese speakers. The most frequent function is found to be feedback by listeners, but speakers also nod a lot in different contexts. It is found, in fact, that the Japanese nod with an average frequency of 5.57 seconds, against 22.5 seconds for Americans. Rehm et al. (2008) have studied differences between German and Japanese speakers' behaviour in first encounters, and found differences in gesture frequency, amplitude and speed, with Germans scoring higher than the Japanese on all three dimensions. Thus, although nods are quite common with Japanese speakers, in general they seem to display less frequent non-verbal behaviour. Allwood and Liu (2010) have analysed cultural differences in first encounter meetings comparing Swedish and Chinese speakers, and found that the Chinese use more laughter, gaze around, gaze sideways and covering their mouth with their hands. On the other hand, only the Swedes use up-down head movements and tilts. Both Swedes and Chinese use more feedback gestures when they speak English in intercultural interactions.

In Hoftede's much quoted theory of cultural dimensions (i.a. Hoftede, 2001), national cultures are placed in a six-dimensional space. The six dimensions are power distance (PDI), individualism vs. collectivism (IDV), acceptance of gender difference (MAS), uncertainty avoidance (UAI), long-term orientation (LTOWVS) and indulgence vs. restraint (IVR). According to this theory (see also Rehm et al. 2008; 2009), at least some of these dimensions can be related to communicative behaviour. For example, high power distance acceptance typically implies more formal behaviour and more physical distance between strangers. Low uncertainty avoidance, on the other hand, goes together with a tendency towards phlegmatic, non-emotional behaviour. In Hofstede's model, the languages targeted in the studies mentioned so far are placed quite differently with respect to the six dimensions. We show five of them in Figure (1).

Clearly, Swedish and Danish cultures are quite similar, and so are Japanese and Chinese, while German culture is more similar to the Scandinavian world in some respects (PDI and IDV) and to the Asian world in others.

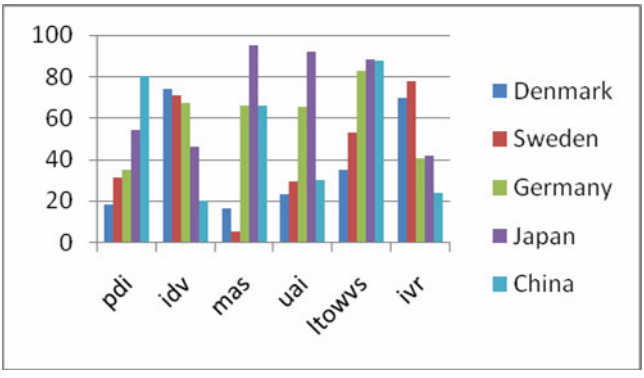


**Fig. 1.** Scoring of different cultures on Hofstede's cultural dimensions (data taken from www.geerthofstede.nl)

## 3 Multimodal Behaviour and Data Collection

Although the differences mentioned in the preceding section cannot univocally be mapped onto expected communicative behaviour patterns in Danish speakers, based on the data we have we can investigate a number of possible systematic relations. The most relevant dimensions in our case are power distance and uncertainty avoidance, which would predict for Danish speakers informal, non-homogeneous and non-emotional behaviour, as opposed to the kind of polite and stereotypical behaviour often ascribed, for instance, to Japanese speakers. To see if these differences emerge in the data, we posed ourselves the following questions:

1. How often do the various types of non-verbal behaviour occur? How uniform or varied is this behaviour? In particular, do Danish speakers nod more or less than speakers of other cultures?
2. Are there significant correlations between the different types and the expression of feedback, and how frequent is multimodal feedback in Danish? How do the results compare with findings for other languages?
3. Do the participants influence each other in their non-verbal behaviour as a sign of politeness?

The Danish corpus consists of 6 videos where pairs (mixed gender or same gender), meet for the first time and talk freely for about 5 minutes. This gives about 30 minutes of conversational interaction. The subjects are standing in a studio, facing each other. The videos were orthographically transcribed, and head movements and facial expressions were annotated by hand in ANVIL (Kipp, 2001) based on the MUMIN annotation scheme (Allwood et al, 1997).

## 4 Analysis of the Data

The total number of gestures identified in the Danish data is 1919. Table (1) shows how they are distributed according to four different gesture shape attributes. Note that these are not mutually exclusive: *Eyebrow*s attributes may occur on their own, but

also in conjunction with a *Face* or a *Head movement* attribute. *Repeat* always occurs together with *Head movement*.

**Table 1.** Facial expressions and head movements in the Danish corpus

| Face (type/#) | | Eyebrows (type/#) | | Head mov. (type/#) | | Repeat (type/#) | |
|---|---|---|---|---|---|---|---|
| Smile | 330 | Frown | 44 | Nod | 249 | Single | 928 |
| Laughter | 143 | Raise | 222 | Jerk | 70 | Repeated | 280 |
| Scowl | 0 | BrowsOther | 1 | HeadBackward | 101 | | |
| FaceOther | 29 | | | HeadForward | 139 | | |
| | | | | Tilt | 214 | | |
| | | | | SideTurn | 182 | | |
| | | | | Shake | 136 | | |
| | | | | Waggle | 31 | | |
| | | | | HeadOther | 86 | | |
| Face total | 502 | Brows total | 267 | Head total | 1208 | Repeat total | 1208 |

Head movements constitute the majority of the gestures, and most of them are single movements. They occur with a frequency of 1.49 seconds, compared with 3.5 for face expressions and 8.7 for eyebrow movements. In relation to the number of words (6000 including filled pauses), there are 0.2 head movements, 0.08 face expressions and 0.03 eyebrow movements per word.

Average and standard deviation for the two most frequent behaviour modalities are given in Table (2).

**Table 2.** Average and SD in the Danish corpus

| | 10 subjects | | 8 subjects | |
|---|---|---|---|---|
| Behaviour | Average | SD | Average | SD |
| Head | 120.8 | 39.46 | 109.37 | 35.58 |
| Face | 43.6 | 18.86 | 36.5 | 13.65 |

The figures in the two leftmost columns refer to the whole population, whilst those to the right have been calculated after the two most deviant subjects were taken out. The amount of standard deviation is not negligible, in that it consists of about 1/3 of each average. The data taken into consideration in this study are a subset of a larger corpus still under development, thus it remains to be seen whether standard deviation will diminish in the larger corpus.

Looking now at more specific behaviours, *Nod* is the most common of the head movements, with a frequency of 7.23. Compared to the figures in Maynard's study, the Danes nod less than the Japanese but still much more than the Americans. *Tilt*, however, is almost as frequent, closely followed by *SideTurn*. In general, finding a

**Table 3.** Feedback and modalities in the Danish corpus

| | Feedback (%) | Other function (%) |
|---|---|---|
| Head | 61 | 64 |
| Face | 28 | 26 |
| Eyebrows | 11 | 11 |
| Total | 100 | 100 |

number of different head movements seems to confirm the expectation of a varied, non-stereotypical non-verbal behaviour in Danish speakers.

The annotation shows that 803 non-verbal behaviours (about 40% of the total 1919) are used to provide feedback, especially to give feedback, but also to a lesser extent to elicit feedback. Looking now at different modalities and feedback, we see (Table 3) that head movements are the most represented type, followed by face expressions and eyebrows attributes. This reflects in part the fact that head movements are in general the most frequent type. In fact, about 60% of the head movements (as opposed to 40% of all movements) is used to express feedback. This is similar to the results obtained by Cerrato (2007) for Swedish. It remains to be seen if a similar analysis on data for other languages would provide interesting differences.

Looking at specific head movement types shows that feedback is especially associated with head nods and smiles, as shown in Table (4).

**Table 4.** Feedback and non-verbal behaviour types in the Danish corpus

| Behaviour | # | % |
|---|---|---|
| FaceOther | 22 | 0.03 |
| Laughter | 61 | 0.08 |
| Smile | 140 | 0.17 |
| HeadBackwards | 48 | 0.06 |
| HeadForward | 51 | 0.06 |
| HeadOther | 24 | 0.03 |
| Jerk | 53 | 0.07 |
| Nod Repeated | 108 | 0.13 |
| Nod Single | 58 | 0.07 |
| Shake | 46 | 0.06 |
| SideTurn | 43 | 0.05 |
| Tilt | 55 | 0.07 |
| Waggle | 6 | 0.01 |
| Frown | 20 | 0.02 |
| Raise | 68 | 0.08 |
| Total | 803 | 1 |

An interesting question is whether a similar analysis carried out on Japanese or Chinese data would show more emotional behaviour, for example a higher number of smiles and laughs as the study in Allwood and Liu (2010) would suggest.

To elucidate the question of whether the participants influence each other in their non-verbal behaviour, we have analysed the multimodal behaviour of two of the subjects (here, A and B) interacting each with two different people. Four videos were thus used in this analysis. The results show that the subjects either kept their non-verbal behaviour more or less unchanged, or for some behaviours produced more or fewer movements in opposition to what done by the interlocutor. Thus in these first encounters, rather than a mirroring effect, there is a slight indication that participants make a larger effort to seem interested and give feedback in cases when the interlocutors are more reserved.

**Table 5.** Head Movements and facial expressions of same participant with two different interlocutors

| Meeting | Speaker | Smile | Laughter | Nod | Shake |
|---|---|---|---|---|---|
| SpeakerA+B | **Speaker A** | 13 | 13 | 11 | 4 |
| | Speaker B | 28 | 3 | 28 | 21 |
| SpeakerA+C | **Speaker A** | 33 | 19 | 8 | 11 |
| | Speaker C | 14 | 6 | 8 | 5 |
| SpeakerD+E | **Speaker D** | 41 | 34 | 12 | 14 |
| | Speaker E | 34 | 16 | 44 | 27 |
| SpeakerD+F | **Speaker D** | 50 | 39 | 25 | 7 |
| | Speaker F | 61 | 5 | 35 | 14 |

In Table 5 we show the figures for the most frequently occurring behaviours in the meetings, which comprise Smile, Laughter, Nod and Shake. The dataset analysed here is far too small to provide any conclusive evidence. However, the provisional answer to the question we started out with is that the participants, if they influence each other at all, do not do it in the sense of copying each other's behaviours.

If it is true that we should expect relatively free, individualistic behaviour from Danish participants, lack of mirroring is probably what we should expect. Again, it would be very interesting to investigate this aspect on larger data material, and to carry out a similar analysis on data from more collectivistic cultures like Japanese or Chinese.

## 5 Conclusion

In this paper we have given an account of a Danish multimodal corpus in which head movements and facial expressions were annotated manually with labels referring to their shape, and analysed with respect to their feedback function. It was shown that the non-verbal behaviour in the corpus confirms expectations based on an understanding of Danish culture according to which Danish speakers are prone to keeping an informal, non-stereotypical and non-emotional style. It was found in fact that

compared to other cultures, which have been described as seeking more stereotypical and polite behaviour, Danes nod less and use more varied non-verbal behaviour in connection with feedback. Still in accordance to what we would expect of a relatively individualistic culture, they do not seem to be subject to mirroring effects in the sense that they do not copy each other's behaviour. This aspect, however, should be analysed in more depth on a larger data collection.

Many other questions can of course be asked on culture-specific multimodal behaviour. For example, specific comparisons on the use of emotional cues such as smiles and laughs could be made. Or the effect of familiarisation could be investigated, for instance by testing whether the non-verbal behaviour changes in the course of the conversations. Finally, differences in the use of feedback words could be studied. In all these cases, cultural dimensions can be used to formulate hypotheses on expected behaviour.

## References

[1] Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C., Paggio, P.: The MUMIN Coding Scheme for the Annotation of Feedback, Turn Management and Sequencing. In: Martin, J.C., et al. (eds.) Multimodal Corpora for Modelling Human Multimodal Behaviour. Special issue of the International Journal of Language Resources and Evaluation, Springer, Heidelberg (2007)

[2] Allwood, J., Lu, J.: Chinese and Swedish multimodal communicative feedback. In: Abstracts of the 5th Conference on Multimodality, Sydney, December 1-3, pp. 19–20 (2010)

[3] Bevacqua, E., Heylen, D., Tellier, M., Pelachaud, C.: Facial feedback signals for ECAs. In: AISB, Annual Convention Workshop "Mindful Environments", Newcastle upon Tyne, UK, pp. 147–153 (2007)

[4] Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., Douville, B., Prevost, S., Stone, M.: Animated conversation: Rule-based generation of facial expression, gesture and spoken intonation for multiple conversational agents. In: Proceedings of SIGGRAPH, Orlando, Florida (1994)

[5] Cerrato, L.: Investigating Communicative Feedback Phenomena across Languages and Modalities. PhD Thesis in Speech and Music Communication, Stockholm, KTH (2007)

[6] Edlund, J., Nordstrand, M.: Turn-taking gestures and hour-glasses in a multi-modal dialogue system. In: Proc of ISCA Workshop Multi-Modal Dialogue in Mobile Environments, Kloster Irsee, Germany (2002)

[7] Duncan, S.: Some signals and rules for taking speaking turns in conversations. Journal of Personality and Social Psychology 23(2), 283–292 (1972)

[8] Duncan Jr., S., Fiske, D.W.: Face-to-Face Interaction: Research, Methods and Theory. Lawrence Erlbaum Associates Publishers, Mahwah (1977); Distributed by John Wiley and Sons

[9] Hadar, U., Steiner, T.J., Grant, E.C., Clifford Rose, F.: The timing of shifts of head postures during conservation. Human Movement Science 3(3), 237–245 (1984)

[10] Hadar, U., Steiner, T.J., Clifford Rose, F.: Head movement during listening turns in conversation. Journal of Nonverbal Behavior 9(4), 214–228 (1985)

[11] Hofstede, G.: Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations. Sage Publications, Thousands Oaks (2001)

[12] Jokinen, K., Navarretta, C., Paggio, P.: Distinguishing the communicative functions of gestures. In: Proceedings of the 5th Joint Workshop on Machine Learning and Multimodal Interaction, Utrecht, The Netherlands, September 8–10, pp. 8–10 (2008)

[13] Jokinen, K., Harada, K., Nishida, M., Yamamoto, S.: Turn-alignment using eye-gaze and speech in conversational interaction. In: Proceedings of Interspeech 2010, Makuhari, Japan (2010)
[14] Jokinen, K., Nishida, M., Yamamoto, S.: Collecting and Annotating Conversational Eye-Gaze Data. In: Workshop on Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality, Proceedings of the Language Resources and Evaluation Conference (LREC 2010), Malta (2010)
[15] Kendon, A.: Gesture: Visible Action as Utterance, Cambridge (2004)
[16] Kipp, M.: Anvil – A Generic Annotation Tool for Multimodal Dialogue. In: Proceedings of Eurospeech 2001, pp. 1367–1370 (2001)
[17] Maynard, S.K.: Interactional functions of a nonverbal sign: Head movement in japanese dyadic casual conversation. Journal of Pragmatics 11(5), 589–606 (1987)
[18] McClave, E.Z.: Linguistic functions of head movements in the context of speech. Journal of Pragmatics 32(7), 855–878 (2000)
[19] Paggio, P., Allwood, J., Ahlsen, E., Jokinen, K., Navarretta, C.: The NOMCO multimodal Nordic resource - goals and characteristics. In: Proceedings of the Language Resources and Evaluation Conference (LREC 2010), Malta (2010)
[20] Pelachaud, C., Poggi, I.: Multimodal Embodied Agents. The Knowledge Engineering Review 17(2), 181–196 (2002)
[21] Rehm, M., Nakano, Y., André, E., Nishida, T.: Culture-Specific First Meeting Encounters between Virtual Agents. In: Prendinger, H., Lester, J.C., Ishizuka, M. (eds.) IVA 2008. LNCS (LNAI), vol. 5208, pp. 223–236. Springer, Heidelberg (2008)
[22] Rehm, M., André, E., Bee, N., Endrass, B., Wissner, M., Nakano, Y., Akhter Lipi, A., Nishida, T., Huang, H.-H.: Creating Standardized Video Recordings of Multimodal Interactions across Cultures. In: Kipp, M., Martin, J.-C., Paggio, P., Heylen, D. (eds.) Multimodal Corpora. LNCS, vol. 5509, pp. 138–159. Springer, Heidelberg (2009)
[23] Yngve, V.: On getting a word in edgewise. Papers from the Sixth Regional Meeting of the Chicago Linguistic Society, p. 568 (1970)