

Feedback and gestural behaviour in a conversational corpus of Danish

Patrizia Paggio

University of Copenhagen
Centre for Language Technology
paggio@hum.ku.dk

Costanza Navarretta

University of Copenhagen
Centre for Language Technology
costanza@hum.ku.dk

Abstract

This paper deals with the way in which feedback is expressed through speech and gestures in the Danish NOMCO corpus of dyadic first encounters. The annotation includes the speech transcription as well as attributes concerning shape and conversational function of head movements and facial expressions. Our analysis of the data shows that all communication modalities, i.e. head, face and eyebrows, contribute to the expressions of feedback, with repeated nods and smiles as the most frequent feedback gesture types. In general, the use of nods as feedback gestures in our data is comparable to what earlier studies have found for other languages, but feedback is also often expressed by other head movements and smiles.

1 Introduction

Head movements and facial expressions play an important function in face-to-face interaction. In particular, many authors have observed that head nods are an important means of expressing what we here call feedback, i.e. unobtrusive behaviour that has the purpose of either giving or eliciting signs of *contact*, *perception*, *understanding* and *agreement* or *disagreement* (Allwood et al., 1992).

Dittmann and Lewellyn (1968), for instance, focus on nodding by listeners, and find that nods occur together with brief feedback responses more often than predicted by chance. Yngve (1970) and Duncan (1972) consider head nods as examples of backchannels, i.e. feedback signals given by the listener without trying to take the floor. Hadar et al. (1985) monitor head movements in five subjects during conversation, and find that agreeing is one of the functions head movements are associated with (the others are wanting to take the turn

and aligning with the interlocutor's stressed syllables and pauses). Maynard (1987) 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. An interesting observation in this study relates to the culture-specificity of gesturing: the Japanese nod with an average frequency of 5.57 seconds (in other words, one nod for every 5.57 seconds), while Americans do so with an average of only 22.5 seconds. McClave (2000), in a qualitative study of head movements in dialogues between two pairs of American speakers, observes that head movements occur together with a whole array of functions and senses, one of which is linked to what she calls backchanneling requests: the speaker nods to ask the listener for feedback, and the listener in turn nods.

Head movements have also been studied in relation to Scandinavian languages, of which Danish, which is targeted in this paper, is an example. It has been observed that 70% of all head movements in a subset of the Swedish GSLC corpus (Nivre et al., 1998) are related to feedback, and that most of these are nods and up-nods (Cerrato, 2007).

While there is a whole body of research on facial expressions as vehicles of emotional response (Hager and Ekman, 1983; Busso and Narayanan, 2007), less attention has been given to the role played by facial expressions with respect to conversational feedback. Smiles and laughter as signals of feedback are studied for instance by Allwood and Lu in this volume and Lu et al. (Under publication), who find that in first encounter situations, both Chinese and Swedish speakers use smiles and chuckles to give feedback.

In previous work (Jokinen et al., 2008), we studied facial expressions and head movements 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. In Paggio and Navarretta (2010) and Navarretta and Paggio (2010) we looked at the relation between head movements and facial expressions on the one hand, and the dialogue act functions of linguistic feedback expressions on the other and showed that head gestures, where they occur, contribute to the semantic interpretation of feedback expressions in a significant way.

Here we present empirical evidence from a multimodal corpus of Danish first encounters of how head movements and facial expressions are used in conversational Danish as signals of feedback giving and eliciting. We start by explaining how the corpus was collected in Section 1. We then describe the annotation categories and procedure used in Section 2. In Section 3 we provide quantitative measures of the annotated data. In Section 4 we briefly discuss how the corpus can be used in machine learning studies of multimodal behaviour and conclude.

2 Corpus collection

The Danish NOMCO corpus is one of a number of multimodal corpora in Swedish, Danish, Finnish and Estonian that have been collected and annotated within the Nordic NOMCO project (Paggio et al., 2010). The aim of the project is to provide comparative annotated multimodal data in the Nordic languages and, based on these data, to investigate how speech and gestures together are used to express feedback, turn taking, sequencing and information structure.

The Danish first encounter corpus consists of 12 dyadic interactions of a duration of approximately 5 minutes each, in which subjects who have not met before try to get to know each other. The participants were six males and six females, all native speakers of Danish aged between 21 and 36, either university students or people with a university education. They did not know each other beforehand, and were not acquainted with the purpose of the recordings. The videos were recorded in the TV studio of the Faculty of Humanities at the University of Copenhagen. The subjects are standing in front of each other and are recorded by three different cameras. The speech is recorded through microphones attached to the ceiling. For each dialogue, two versions were produced, one showing a

long shot of the two participants facing each other, the other combining two mid shots taken from different angles into a split video. The two views are shown in Figure 1.

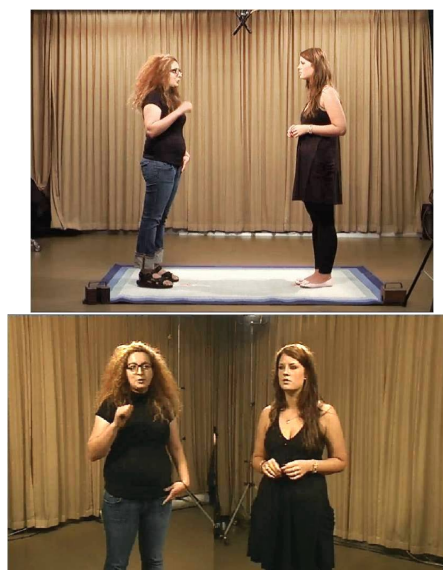


Figure 1: Recordings from the Danish NOMCO dialogues: total and split views

A questionnaire was given to the participants to collect information on how they experienced the conversations. They were asked to rate their experience along a number of parameters concerning their emotional state and the interaction itself. The results indicate that the subjects were not too affected by the artificial setting even though they were aware of it. In particular, since the scores for perturbedness, tenseness and awkwardness were all below average, we consider the corpus a relatively valid exemplification of natural interaction. For a more detailed analysis of the questionnaire results, see Paggio and Diderichsen (2010).

3 Annotation categories and procedure

3.1 Orthographic transcription

The first step in the annotation process was to produce an orthographic transcription of the audio signal. This was done using Praat (Boersma and Weenink, 2009). The transcription includes word boundaries as well as word stress, indicated by a “,” before the stressed vowel. Pauses are represented by a “+”, and filled pauses glossed with English words, e.g. *laugh*, *breath* or expressions such as *øh*. The Praat transcriptions were then imported into the ANVIL tool (Kipp, 2004), which was

used for gesture annotation. In ANVIL, sentence boundaries, in the front of an attribute *boundary true*, were added to the transcription based on the occurrence of pauses as well as on syntactic criteria. Furthermore, topic and focus were identified in each sentence, and the attributes *topic true* and *focus true* were added to the corresponding words according to the methodology described in (Paggio, 2006a; Paggio, 2006b). In short, *topic* indicates the presupposed entity about which the sentence predicates something new, while *focus* indicates non-presupposed information.

Word token	topic	focus	boundary
+	false	false	<i>true</i>
jeg	<i>true</i>	false	false
hedder	false	<i>true</i>	false
H,anne	false	<i>true</i>	false
+	false	false	<i>true</i>

Table 1: Topic and focus annotation example

In Table 1 we show in table format the assignment of topic, focus and clause boundary attributes to the utterance *jeg hedder Hanne* (lit: I call Hanne, or “My name is Hanne”) from one of the NOMCO dialogues. Boundaries are placed together with the pauses that precede and follow the sentence, *jeg* (I) refers to topic, i.e. the entity about which the sentence predicates something new, whilst *hedder Hanne* (lit: call Hanne), which contains the only stressed word, is the focus, i.e. the new information.

3.2 Gesture annotation scheme

Gestures in the NOMCO data are annotated with a subset of the attributes defined in the MUMIN annotation scheme (Allwood et al., 2007). The MUMIN scheme is a general framework for the study of gestures in interpersonal communication that has been applied to multimodal data in several languages within the context of the Nordic MUMIN network (www.cst.dk/mumin). It concerns facial expressions, head movements, hand gestures and body posture, and it provides attributes for shape as well as function.

The attributes for the annotation of gesture shape used in this study are shown in Table 2. The granularity of the annotation categories is deliberately coarse in that we only want to be able to distinguish different communicative functions rather than provide precise morphological descriptions.

The functional annotation features in MUMIN concern feedback, turn management and sequenc-

Modality	Attribute	Value
Head	HeadMovement	Nod, Jerk, HeadForward, HeadBackward, Tilt, SideTurn, Shake Waggle, HeadOther
	HeadRepetition	Single, Repeated
Face	GeneralFace	Smile, Laugh, Scowl, FaceOther
	Eyebrows	Frown, Raise, BrowsOther

Table 2: Shape Annotation Features for Head and Face

Attribute	Value
Basic	ContactPerceptionUnderstanding (CPU), BasicOther
Direction	FbGive, FbElicit, FbGiveElicit, FbUnderspecified
Agreement	Agree, NonAgree

Table 3: Functional annotation of feedback gestures

ing. In this study, however, only feedback attributes will be considered. They are shown in Table 3.

The *Basic* attribute has two possible values: *ContactPerceptionUnderstanding (CPU)* indicates that participants are willing and capable of interacting, perceiving and understanding what is being communicated (Allwood et al., 1992); *BasicOther* is used if one of the above dimensions, e.g. understanding, appears to be lacking (this does not occur in the current corpus, thus only CPU is used) If *Basic* is coded, a value for the *Direction* attribute has to be chosen, too. We distinguish between i. *FeedbackGive*, where the listener gives feedback (often called backchannelling), ii. *FeedbackElicit*, where the speaker appears to be eliciting feedback from the listener, iii. a combination of both values, and iv. an underspecified value. Finally, a feedback gesture may express agreement or disagreement towards a statement, for which the scheme foresees the two values *Agree* and *NonAgree*.

In addition to the shape and function attributes, for each gesture a relation with the corresponding speech expression, if one such exists, is also annotated by means of a link. The link can point to a speech segment uttered by the person producing the gesture (by means of the attribute *MMRelationSelf*), or to a speech segment in the interlocutor’s vocal stream (by means of the attribute *MMRelationOther*).

3.3 Gesture annotation procedure

Three annotators, all of them students of linguistics, created the annotation. To ensure reliability, they received an initial training where they all worked together coding the same video. Then a second video was coded by each of them separately. The results were discussed and corrected, and a set of written guidelines were developed based on these discussions. In this preliminary exercise, the *Cohen's kappa* (Cohen, 1960) figures obtained were on average for the three pairs of coders in the range 0.5-0.6 for face attributes and 0.6-0.8 for head movements. Considering the fact that the agreement measure calculated in ANVIL reflects agreement of segmentation as well as labelling, these figures are quite satisfactory.

Each of the remaining videos was subsequently annotated by one of the coders and corrected by the other. Disagreements were again discussed and evened out. If the two coders still could not agree, a third annotator made the final decision. Throughout this process, the guidelines were continually improved with examples and explanations. After having annotated five videos following this procedure, we repeated the inter-coder agreement exercise between the two annotators who had shown most disagreement the first time, and noted an improvement of about 10% for both face and head gestures.

To annotate facial expressions and head movements according to this procedure takes on average 2 hours per minute per speaker including discussions and subsequent corrections.

4 Data analysis

So far, nine of the twelve videos have been annotated and analysed. The total duration of this annotated material is 3027 seconds, in other words 50 minutes and 45 seconds. The length of the individual annotated clips varies from about 140 seconds to about 360. The total number of word tokens (including filled pauses) is 10800. The total number of gestures identified is 3391.

4.1 Gesture frequency

Table 4 shows how gestures are distributed according to the three major shape attributes. Note that the *Eyebrows* gestures listed here are those occurring without a concomitant general facial expression like *Smile* or *Laughter*. Head movements are also coded with a value for repetition. The dis-

FaceGeneral		Eyebrows		HeadMovement	
Smile	499	Raise	263	Nod	520
Laughter	198	Frown	85	Tilt	388
FOther	45	BOther	3	SideTurn	328
Scowl	5			HForward	264
				Shake	257
				HBackward	200
				HOther	148
				Jerk	122
				Waggle	66
Face total	747	Brows total	351	Head total	2293

Table 4: Gesture types in the Danish NOMCO corpus

tribution is 1714 *Single* movements and 579 *Repeated* ones. Head movements constitute the majority of the gestures, and most of them are single movements.

Type	No	sec/g	g/w	g/sec
All gestures	3389	0.89	0.31	1.12
Head	2291	1.32	0.21	0.76
Nods	520	5.82	0.05	0.17
Face	747	4.05	0.07	0.25
Eyebrows	351	8.62	0.03	0.12

Table 5: Gesture type frequency

In Table 5 we show the frequency counts for some of the most frequent gesture types. The second column shows the raw counts, the third one the proportion of seconds per gesture, the fourth one the proportion of gestures per word, and the last one the proportion of gestures per second. The proportion of seconds per gesture allows us to compare with the findings in the already mentioned study by Maynard, where it is claimed that Japanese speakers make a nod every 5.5 seconds. The figure for Danes is one nod every 5.6 seconds, which is very similar. This seems to show that Danes and Japanese behave similarly as far as nodding is concerned - at least in the sense that they nod with similar frequencies. However, the subjects in Maynard's study already knew each other, so the datasets are not directly comparable. Moreover, we have not looked at dimensions concerning the amplitude or velocity of the nods, where differences may indeed arise. A discussion of how differences in gestural behaviour can be couched in the perspective of cultural diversity can be found in Paggio and Navarretta (2011).

An interesting issue is how much individual difference can be observed in a corpus which is trying to model culture-specific behaviour in a certain communication situation, or activity. Table 6

Modality	Average No	SD
Face	61.00	26.80
Head	127.28	34.61

Table 6: Number of facial expressions and head movements: average and standard deviation

shows the average number of facial expressions (this time including eyebrows) and head movements together with standard deviation figures. The variation is especially large in the case of facial expressions, suggesting that one should be cautious in generalising from these data, and that more data should be added to the corpus to provide a more reliable basis for quantitative studies of facial expressions. A question that we will investigate further in these data is whether the deviation in gesture production is dependent on the amount of speech produced by the gesturer and by the interlocutor.

4.2 Gesture and feedback

Out of the total 3391 gestures identified in the corpus, 1594 (47%) have been annotated with the *Basic CPU* feedback feature. This means that on average, there is a feedback gesture either by the speaker or by the interlocutor every 0.3 seconds. Is this what one should expect? In order to answer the question, it may be useful to compare with other corpora in Danish or similar corpora in other languages.

Corpus	No g	g/w	FB g	No w	FB w
NOMCO	3391	0.3	47%	10,800	0.06%
DanPASS	264	0.05	21%	5,556	7.00%

Table 7: Feedback in the NOMCO and DanPASS corpora

We can start by looking at feedback in the DanPASS dialogues, which are part of a corpus of spoken Danish (Grønnum, 2006) in which two speakers have to solve a map-task. The subjects sit in separate studies without being able to see each other, and they talk through headsets. Given the very different settings as well as the different genres (map-oriented dialogue vs free conversation), we would expect more feedback words (*yes*, *no*, and similar) and less feedback gestures in DanPASS as opposed to NOMCO. We have used a small sub-set of this corpus (8 videos) for earlier studies, where head movements and facial expressions were annotated following the same method-

ology as in NOMCO. In Table 7 we show how this sub-corpus compares with the NOMCO data on a number of parameters. As expected, the number of gestures by word is in general much lower in DanPASS, and the proportion of gestures that are used for feedback is also lower. We have not conducted an analysis of the functions of the remaining gestures, we can only guess that they may have a turn taking or focusing function. Finally, the percentage of feedback words is as expected much higher in DanPASS compared to NOMCO. Participants in a task-oriented dialogue that cannot see each other need to check mutual understanding and grounding by using feedback words.

Gesture	No	%
Nod Repeated	250	0.16
Smile	248	0.16
Nod Single	134	0.08
Tilt	125	0.08
Raise	117	0.07
Shake	112	0.07
HeadBackward	110	0.07
HeadForward	99	0.06
Jerk	92	0.06
Laughter	91	0.06
SideTurn	84	0.05
Frown	40	0.03
HeadOther	40	0.03
FaceOther	32	0.01
Waggle	20	0.01
Total	1594	1

Table 8: Feedback distribution in the Danish NOMCO corpus

While it is easy to see that NOMCO is different from a map-task dialogue with respect to gestural behaviour in general, and to gestural feedback in particular, it is not so straightforward to compare it with similar corpora in different languages. The NOMCO project is working on a comparison between Danish, Swedish and Finnish data. Here, we will hold the Danish NOMCO data against earlier findings on the use of nods as feedback signals in Japanese and Swedish and Japanese.

Table 8 shows how feedback gestures in the Danish NOMCO corpus are distributed among different gesture types. Head movements are in general the preferred feedback modality. In fact, about 67% of the head movements (as opposed to 47% of all movements) is used to express feedback. This is similar to the results obtained by Cerrato (op.cit.) for Swedish. If we look at specific movement types, nods are by far the most common type. We have seen that nods occur roughly as often in our corpus as in the Japanese data studied by May-

nard (op.cit.), i.e. every 5-6 seconds. In the Danish data, in 54.61% of the cases, nods are used to express feedback. In the Japanese data, Maynard claims that nods are used as feedback signals in almost 50% of the cases (other functions mentioned in this study are turn shifts, emphasis and clause boundary marking). Thus, the Danish and Japanese data also seem similar on this dimension, although, as already pointed out, these comparisons should be taken with due caution because not all aspects are kept equal in the two corpora.

In general we can conclude that the use of head movements and facial expressions as feedback signals in the NOMCO corpus confirms earlier findings concerning the pervasiveness of the phenomenon as well as the frequent use of nods as feedback signals. However, our data also show that other head movements, such as tilts, shakes and head-backward movements, are often used to express feedback. Finally, to conclude this section, the data also allow us to see which of the feedback directions is the most frequent. In 77% of the cases feedback is given, in 20% it is elicited, and in 3% of the cases both directions seem present at the same time.

5 Conclusion

The analysis of feedback in a multimodally annotated Danish corpus of first encounters shows that both speech and gestures (in the present study head movements and facial expressions) are used, alone or in combination, to give and elicit feedback. The most frequently used feedback-related gestures in the data are head movements, especially repeated and single nods, confirming preceding studies of multimodal feedback. However in our corpus also other types of head movement and various facial expressions have been recognised to have a feedback-related function.

Comparing feedback expressions in this corpus and in a map-task corpus we found that feedback was expressed more frequently with gestures in the former, and verbally in the latter. These results are not surprising given the nature and the settings of the two corpora.

The analysis of the annotated data also indicates that there is a large individual variation in the frequency with which the interaction participants used gestures to express feedback. This is especially true for facial expressions. In future we will investigate the relation between individ-

ual frequency of speech and gesture production in the NOMCO data. Furthermore, future work still related to the study of feedback will also comprise the comparison of feedback expression in first encounters corpora in two other Scandinavian languages for which these corpora have been collected and annotated.

While the focus of this study has been on gestural feedback, the Danish NOMCO corpus of first encounters provides the means to investigate the interaction of speech and gestures with respect to a number of conversational functions, especially turn taking and information structure. The rich functional annotation of gestures will be analysed against the focus and topic tags but also in comparison with automatically extracted prosody features. Finally, we also plan to annotate hand gestures to provide a comprehensive analysis of the multimodal behaviour in the corpus.

Acknowledgements

We would like to acknowledge our partners in the NOMCO project Elisabeth Ahlsén, Jens Allwood and Kristiina Jokinen, as well as the annotators Sara Andersen, Josephine B. Arrild, Anette Studsgård and Bjørn Wesseltolvig. The NOMCO project, the full name of which is “Multimodal Corpus Analysis in the Nordic Countries”, is funded by the NOS-HS NORDCORP programme, see <http://sskkii.gu.se/nomco/>. The work of the Danish group in the project is also funded by the Danish Research Council for the Humanities, see <http://cst.dk/vkk/uk/>.

References

- Jens Allwood, Joakim Nivre, and Elisabeth Ahlsén. 1992. On the semantics and pragmatics of linguistic feedback. *Journal of Semantics*, 9:1–26.
- Paul Boersma and David Weenink. 2009. Praat: doing phonetics by computer (version 5.1.05) [computer program]. Retrieved May 1, 2009, from <http://www.praat.org/>.
- Carlos Busso and Shrikanth S. Narayanan. 2007. Interrelation between speech and facial gestures in emotional utterances: A single subject study. In *IEEE Transactions on Audio, Speech and Language Processing*, volume 15, pages 2331–2347, July.
- Loredana Cerrato. 2007. *Investigating Communicative Feedback Phenomena across Languages and Modalities*. Ph.D. thesis, Stockholm, KTH, Speech and Music Communication.

- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.
- Allen Dittmann and Lynn Llewellyn. 1968. Relationship between vocalizations and head nods as listener responses. *Journal of Personality and Social Psychology*, 9.
- Starkey Duncan. 1972. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23:283–292.
- Nina Grønnum. 2006. DanPASS - a Danish phonetically annotated spontaneous speech corpus. In Nicoletta Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk, and D. Tapias, editors, *Proceedings of the 5th LREC*, pages 1578–1583, Genoa, May.
- Uri Hadar, T. J. Steiner, and F. Clifford Rose. 1985. Head movement during listening turns in conversation. *Journal of Nonverbal Behavior*, 9(4):214–228, December.
- Joseph C. Hager and Paul Ekman. 1983. The Inner and Outer Meanings of Facial Expressions. In J. T. Cacioppo and R. E. Petty, editors, *Social Psychophysiology: A Sourcebook*, chapter 10. The Guilford Press, New York.
- Kristiina Jokinen, Costanza Navarretta, and Patrizia Paggio. 2008. Distinguishing the communicative functions of gestures. In *Proceedings of the 5th MLMI*, LNCS 5237, pages 38–49, Utrecht, September. Springer.
- Michael Kipp. 2004. *Gesture Generation by Imitation – From Human Behavior to Computer Character Animation*. Boca Raton, Florida: Dissertation.com.
- Jia Lu, Jens Allwood, and Elisabeth Ahlsén. Under publication. A study on cultural variations of smile based on empirical recordings of Chinese and Swedish first encounters. In Dirk Heylen, Michael Kipp, and Patrizia Paggio, editors, *Proceedings of the workshop on Multimodal Corpora at ICMI-MLMI 2011*, Alicante, Spain, Nov.
- Senko Maynard. 1987. Interactional functions of a nonverbal sign: Head movement in Japanese dyadic casual conversation. *Journal of Pragmatics*, 11:589–606.
- Evelyn McClave. 2000. Linguistic functions of head movements in the context of speech. *Journal of Pragmatics*, 32:855–878.
- Costanza Navarretta and Patrizia Paggio. 2010. Classification of feedback expressions in multimodal data. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, pages 318–324, Uppsala, Sweden, Juli 11-16.
- Joakim Nivre, Jens Allwood, Jenny Holm, Dario Lopez-Kästen, Kristina Tullgren, Elisabeth Ahlsén, Leif Grönqvist, and Silvana Sofkova. 1998. Towards Multimodal Spoken Language Corpora: TransTool and SyncTool. In *Proceedings of the Workshop on Partially Automated Techniques for Transcribing Naturally Occurring Speech at COLING-ACL '98*, Montreal, Canada, August.
- Patrizia Paggio and Philip Diderichsen. 2010. Information structure and communicative functions in spoken and multimodal data. In Peter Juel Henriksen, editor, *Linguistic Theory and Raw Sound*, volume 49 of *Copenhagen Studies in Language*, pages 149–168. Samfundslitteratur.
- Patrizia Paggio and Costanza Navarretta. 2010. Feedback in head gesture and speech. In Kipp et al., editor, *Proceedings of LREC-2010*, pages 1–4, Malta, May 17.
- Patrizia Paggio and Costanza Navarretta. 2011. Head movements, facial expressions and feedback in Danish first encounters interactions: a culture-specific analysis. In C. Stephanidis, editor, *Universal Access in Human-Computer Interaction. Users Diversity. Proceedings of 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011*, pages 583–590, Orlando, FL, USA, July. Springer.
- Patrizia Paggio, Jens Allwood, Elisabeth Ahlsén, Kristiina Jokinen, and Costanza Navarretta. 2010. The NOMCO multimodal nordic resource - goals and characteristics. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).
- Patrizia Paggio. 2006a. Annotating information structure in a corpus of spoken Danish. In *Proceedings of the 5th International Conference on Language Resources and Evaluation LREC2006*, pages 1606–1609, Genova, Italy.
- Patrizia Paggio. 2006b. Information structure and pauses in a corpus of spoken Danish. In *Conference Companion of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, pages 191–194, Trento, Italy.
- Victor Yngve. 1970. On getting a word in edgewise. In *Papers from the sixth regional meeting of the Chicago Linguistic Society*, pages 567–578.