

# Challenges Ahead

## Head movements and other social acts in conversations

Dirk Heylen\*

\*University of Twente  
PoBOX 217, 7500 AE Enschede, The Netherlands  
heylen@ewi.utwente.nl

### Abstract

When involved in face-to-face conversations, people move their heads in typical ways. The pattern of head gestures and their function in conversation has been studied in various disciplines. Many factors are involved in determining the exact patterns that occur in conversation. These can be explained by considering some of the basic properties of face-to-face interactions. The fact that conversations are a type of joint activity involving social actions together with a few other properties, such as the need for grounding, can explain the variety in functions that are served by the multitude of movements that people display during conversations.

## 1 Introduction

People involved in face-to-face conversations move their heads in typical ways. We want to know more about the kinds of movements and movement patterns that occur and about the factors that determine these. Who would disagree that on the whole the pattern of head movements people display in conversations seems to differ significantly from the patterns found in non-conversational settings; when people are alone, for instance? Although this may appear to obvious to be worth stating, it is not totally insignificant. Because it clearly suggests that one can assume that the primary determinants of these particular displays have to do with the nature, the purpose and the organization of face-to-face conversations.

The call for papers to the Social Cues workshop raises the following issue: “Animals that live in same species groups, including humans, develop protocols for dealing with intra group pressures. These protocols require the presentation and recognition of cues that express social relations and any agent, human or virtual, that is to operate in a social context must be able to work with these cues. A key question is what protocols and techniques have evolved in human society, and what must an Embodied Conversational Agent do to be a recognisably social being?”

Embodied conversational agents are designed to take part in face-to-face conversations with humans. The answer to the second question could therefore simply be: the agent should know how to engage

in a face-to-face conversation. Properly engaging oneself in a conversation entails having internalized how to deal with the protocols and techniques that have evolved in human society and knowing how to turn the result into linguistic action. Using language is a form of social action. To find out what the protocols and techniques are, one can simply turn to all the research literature on what is involved in having a conversation. This subject has been studied by many research traditions, including anthropology, sociology, social psychology, ethology, personality psychology, psychiatry, linguistics, anthropological linguistics, cognitive psychology, philosophy, ethnomethodology, micro-sociology, neuropsychology and psycholinguistics (Duncan and Fiske, 1977). Creators of virtual humanoids can or should incorporate in their design and implementation everything that is known about what face-to-face conversation involves. Or they can pursue further studies along these lines for those behaviours that have not been sufficiently analysed in the literature to incorporate in the computational models.

By what systems of “rules” or “conventions” are face-to-face conversations organized? The interest in conversations shown by the various disciplines is evidence for the many levels on which organizational rules are defined: linguistic conventions (related to lexical issues, syntax and semantics), conversation conventions (programs or scripts on how to enter and exit conversations, to take turns) task and specific domain conventions, and social conventions (knowing

what is appropriate action). These do not function independently. For instance, the rules that regulate turn taking (conversational conventions) also involve social parameters. Consider the issue when is it appropriate to interrupt. In this case, various aspects of the way people relate to each other, their status, dominance and other factors play a role in whether or not (and how) this is done. Another example of how the levels connect has to do with how acts on one level are made up of acts on another level. Specific conversational tasks (negotiation, teaching) involve a specific sequence of dialogue acts on a lower level.

An important challenge for the research on embodied conversational agents is how to integrate these ideas, observations, and theories from the various disciplines and how to put them into rules and procedures that embodied agents can use in actual interaction. Embodied Conversational Agents research has always been a highly eclectic business. ECA researchers borrow insights from linguistics, cognitive science, AI, cognitive psychology and social psychology. The social perspective has become increasingly important in the work on embodied conversational agents, witness this workshop but also work on friendship and long term relations with ECA's (Bickmore, 2003), on social rapport (Bickmore and Cassell, 2005), engagement (Sidner et al., 2004) and the incorporation of politeness theory in the design of a tutor agent (Johnson et al., 2004), for instance. A similar trend is visible in our own work on socially intelligent agents where we have moved from implementing an embodied version of a task-oriented spoken dialogue systems (Nijholt and Heylen, 2002) to the design of socially intelligent agents (Heylen et al., 2004). This involves a shift in perspective. Increasingly we have come to view language as social action. Behaviours of agents are not only designed for their communicative functions (providing information on the task, regulating conversational flow) but the conversation is part of a social encounter. For instance, in building an Intelligent Tutoring System (INES), (Heylen et al., 2004), we made an effort to define dialogue acts using social variables. A tutor has to steer and motivate the student, know when the student welcomes a hint, etcetera. The emotional state related to this form of social interaction typically involves elements and variables such as: social rewards, dependence, status, power, and face. In general one of the goals that people want to come out of social interaction is to enhance the self of each actor (see Argyle (1969)). In the INES case, we therefore decided to incorporate the social variables into our choice of speech act primitives.

In this paper, we take a look at a particular kind of behaviour that people display in face-to-face conversations: head movements. How and why do people move their heads? We will first survey some of the literature that has been devoted to these questions. This will show the many factors involved. A more systematic view arises when we look at the survey from the perspective of a single framework, a view on language as social action articulated in the work by Clark (1996). This provides a way to integrate multiple perspectives on the protocols and techniques that people use in face-to-face interactions.

## 2 Head Movements

The subject of head movements during conversations has been discussed by several researchers from various disciplines, though compared to the studies on gestures and facial expressions, head movements have received far less attention. We consider first, the way the movements as such have been analysed and described to find out what properties of the movements can play a function in the face-to-face encounters. Next we consider the various functions that have been ascribed to these movements.

### 2.1 The Movements

Although it is not the major objective of this paper to look at the properties of head movements as such, it still seems appropriate to outline the various dimensions along which movements can be distinguished.

Ray Birdwhistle who devised several coding schemes for all kinds of kinetic behaviours distinguishes the following head movements: (1) a full nod up and down or down and up, (2) a half nod either up or down, (3) a small "bounce" at the end of (1) or (2), (4) a full side and back sweep (which may contain a nod or half nod) and (5) a cocked head (Birdwhistle, 1970).

The conversational character RUTH, (DiCarlo et al., 2004), allows the same general head movements. The head can nod up and down, rotate horizontally left and right and tilt at the neck from side to side. Furthermore, it can bring the whole head forward or backward.

In Iwano et al. (1996), who analyzed the head movements in a natural dialogue and movements during a cooperative problem solving task, movements were classified in whether they were horizontal, vertical or inclined and whether they were large or small. Combinations of Inclination-Vertical and Inclination Horizontal were also noted.

D	Nods downward
U	Nods upward
F	Brings the whole head forward
B	Brings the whole head backward
R	Turns to model's right
L	Turns to model's left
J	Tilts whole head counterclockwise (around nose)
DR	Nods downward with some rightward movement
UR	Nods upward with some rightward movement
DL	Nods downward with some leftward movement
UL	Nods upward with some leftward movement
TL	Tilts clockwise with downward nodding
TR	Tilts counterclockwise with downward nodding

These classifications only distinguish between different head positions. But as we are talking about head movements we are interested in the changes in head position over time. So there are still other features of movements that may be significant. Hadar et al. (1983b) for instance, also looked at different properties of the movement such as velocity and amplitude. Also Smid et al. (2004) take into account the speed with which certain movements are executed.

Head orientations, speed and amplitude of movements are all basic features of the movement that play a role in distinguishing between different types of movements and they may each contribute in their own way how a movement is interpreted.

An important question is how to segment the movements into significant units. Typically, nods and sweeps are movement patterns that are considered to be significant units in this respect. Graf et al. (2002) found the following typical patterns in their corpus: (1) nod: an abrupt swing of the head with a similarly abrupt motion back; (2) nod with an overshoot at the return (looks like an 'S' lying on its side) and (3) an abrupt swing of the head without the back motion. When looking at syntagmatic relations, other properties may become important. Hadar et al. (1985) also looked at the cyclicity of head nods and shakes of listeners with respect to their difference in communicative function.

The timing with respect to other signals may also bear significance. Several authors (see below) have looked at the relation between head movements and speech. Also the relation between head movement and facial expressions are of interest. In the discussion of the functions of head movements we will occasionally refer to such properties.

## 2.2 The Functions

With the head movements in Table 1, DiCarlo et al. (2004) associate a rough list of functions.

D	General indicator of emphasis
U	indicates a "wider perspective"?
F	indicates the need for "a closer look"?
B	emblem of being "taken aback"?
R	indicates there is more information?
L	indicates there is more information?
J	indicates expectation of engagement from partner?
DR	combines meaning of D and R
UR	combines meaning of U and R
DL	combines meaning of D and L
UL	combines meaning of U and L
TL	indicates contrast of related topics
TR	Perhaps indicates contrast of related topics

Based on the literature on head movements one can put together quite an extensive list of functions and determinants of head movements during conversations<sup>1</sup>. Head movements can have the function to (1) signal yes or no, interest or impatience, (2) enhance communicative attention, (3) anticipate an attempt to capture the floor, (4) signal the intention to continue, (5) express inclusivity and intensification, (6) control and organize the interaction, (7) mark the listing or presenting of alternatives, (8) mark the contrast with the immediately preceding utterances. Furthermore, synchrony of movements may (9) communicate the degree of understanding, agreement, or support that a listener is experiencing. Greater activity by the interviewer (e.g., head nodding) (10) indicates that the interviewer is more interested in, or more emphatic toward, the interviewee, or that he otherwise values the interviewee more. Head movements serve as (11) accompaniments of the rhythmic aspects of speech and typical head movement patterns can be observed marking (12) uncertain statements and (13) lexical repairs. Postural shifts (14) mark switches between direct and indirect discourse.

Considering that the gaze behavior of people in conversations might also involve movements of the head it is also worth considering the functions of gaze and the avoidance of eye-contact in conversations as determinants of head movements. Gaze behaviour has been observed to play a role in (14) indicating addresseehood, (15) effecting turn transitions, (16)

<sup>1</sup>Source include: Argyle (1969), Argyle and Cook (1976), Bernieri and Rosenthal (1991), Dittmann (1972), Goodwin (1981), Freedman (1972), Hadar et al. (1983a), Hadar et al. (1985), Heritage (1989), Kendon (1972), McClave (2000).

the display of attentiveness. When doing (17) a word search a typical gaze pattern occurs. Gaze may (18) reflect the social status. Looking away (19) is used to avoid distraction, to concentrate, (20) to indicate one doesn't want to be interrupted. One looks to the other in order (21) to get cues about mood and disposition of the other, (22) to establish or maintain social contact. Gazing away (23) may reflect hesitation, embarrassment or shyness. Gaze is used (24) to locate referents in abstract space, (25) as backchannel requests, etcetera.

What this list shows is that simple behaviours such as head movements can have many functions and are determined by many variables. The actions may have a clear semantic value, may find their use in managing the conversational process, be expressive of the mental state of the speaker or hearer (their mood, emotions, personality, or cognitive processing) and relate to interpersonal goals and attitudes. We will briefly go into a few of these in the following paragraphs (adapting the classification of movements made in McClave (2000)). In Section 3 we make an attempt to map these functions within a general framework that takes language to be a form of joint, social action.

**Head movements and Speech** The relation between head movements and speech has been investigated in many papers. Dittmann (1972) and Kendon (1972), provided many observations, related to the timing of certain head movements with respect to the speech. One of the observations by Dittmann (1972) was, for instance "that there is a 'significant' but not very close relationship between speech rhythm and body movement. Both hesitations in speech and body movements tend to appear early in phonemic clauses and, in addition, movements tend to follow hesitations wherever they may appear in clauses." Postural shifts of the head are claimed to indicate encoding difficulties.

Hadar and colleagues (see the works cited earlier) have also studied motoric functions of head movements during speech. Parallel to the relationship of hand gestures to speech, it appears that the head moves almost constantly during speech whereas it remains mostly motionless during pauses and while listening. They also found a correlation between head movements and loudness of the speech: "rapid head movements were accompanied by primary peaks of loudness". As a large proportion of head movements is synchronised with speech features such as loudness or pitch, they can be seen as prosody markers in the visual domain (see Graf et al. (2002), for example). In this way they serve similar functions - to mark promi-

nence, for instance.

Bernieri and Rosenthal (1991) write "The astonishing finding in the literature, however, is not that our body is synchronized with our verbal utterances but that our body tends also to coordinate with the verbal utterances of anyone we happen to be listening to at the time." According to Hadar et al. (1985), approximately one fourth of all head movements by listeners occur synchronously with the speaker's speech (see the authors cited, for further references).

**Conversation Management** As the list of functions of head movements and gaze above shows, head movements seem to play an important role in managing the interaction, i.e. in turn-taking and backchanneling processes. McClave (2000) notes that "the 'speech-preparatory' repositioning of the head before the start of talk can simultaneously signal the assumption of a turn or the intention to continue and a such is a part of conversational management." Hadar et al. (1983b) determined that postural shifts co-occurred most significantly between sentences or clauses that were associated with assuming or yielding a turn (see also Duncan (1972)).

Many backchannels by hearers are responses to speakers' nonverbal requests for feedback in the form of up-and-down nods. Listeners recognize and respond to these requests in a fraction of a second.

**Discourse functions** Kendon (1972) notes that the particular patterns of movement vary according to the discourse function of the utterance. For example, in his corpus the speaker's head position during a parenthetical remark contrasted with that during statements that "move the substance of the discourse forward" (o.c., p. 193). Kendon finds a recurrent pattern for most locutions made by the subject who's behaviour he is studying. "At the beginning of each of X's locutions, the head is held either erect and central, or it is held erect and cocked somewhat to the right. As the locution ends, the head is tilted forward or lowered and, in several cases, it is either turned or cocked to the left." The exceptions to this pattern, Kendon argues, have to do with a different discourse function of the locutions. "Of the exceptions, locutions 14 and 16 are parenthetical insertions, locution 4 represents a locution begun again as a correction for locution 3. In this case, it ends with a lowered head. Locution 1 is a 'temporizer' or 'floor acceptance' signal".

Related to such markers of discourse function, McClave groups several functions of head movements as "narrative". The first function is that of marking switches from indirect to direct discourse, marked

with a new orientation of the head. The second function concerns the expression of mental images of characters. An example from her corpus was someone moving her head downward iconically when quoting someone talking to someone smaller. These functions mark the status or function of a discourse fragment. The third function McClave categorizes as “narrative” is deictic and concerns the referential use of space. She also notes a typical kinetic pattern when items in a list or alternatives are presented. “Characteristically, the head moves with each succeeding item - often to a contrasting position”.

**Cognitive processing** When a speaker utters a word or words and immediately rejects this as inappropriate and repairs, the repair is typically preceded or accompanied by head movements (most common: lateral shakes, often small lateral tremors). Above we already indicated that hesitations are often accompanied by head movements. The “thinking face”, described in Goodwin and Goodwin (1986), which involves a turn away from the addressee and a distant look in the face is a stereotypical expression to signal thinking.

**Propositional** Some head movements have a symbolic meaning. Nods are used to signal affirmation and head shakes signal negation in many cultures. McClave (2000) points out that head movements can also express other semantic concepts such as intensification and inclusivity. Intensification is conveyed by head shakes and lateral movements co-occurring with words such as “very”, “a lot”, etcetera. These are considered by Goodwin and Goodwin (1986) as prototypical assessment markers. Inclusivity is expressed by a lateral sweep co-occurring with concepts of inclusivity with words such as “everyone” or “anything”. Uncertainty, marked verbally by phrases such as “I guess”, “I think”, etcetera, are kinesically marked by “lateral shakes whose trajectories may be quite contained”.

If one compares this list of functions and determinants of head movements to those that have been assigned with gaze patterns, one can easily show some overlap. In part this is self-evident, because shifts in gaze often involve shifts in head orientation.

**Gaze** Gaze has various functions in social interaction. Head movements may result from an attempt to gaze towards an interlocutor or away or an attempt to obtain gaze. In this way, the various factors that determine gaze behaviour may also be responsible for changes in head-orientation. Argyle and Cook (1976)

contains an extensive description of the functions of gaze.

1. Speakers look to obtain immediate feedback on the reactions of listeners.
2. Listeners look to supplement auditory information by visual cues
3. Gaze is involved in signalling interpersonal attitudes (people look more at those they like, people high in dominance look more in competitive situations, people high in affiliative needs look more in a cooperative situation, negative attitudes may be signalled by looking away)
4. Shifts of gaze are systematically coordinated with the timing of speech, and help with synchronizing. This is related to the interactional function of the head movements.
5. Gaze is said to be a cue for intimacy.
6. Speakers tend to look away to avoid distraction - particularly at the planning face of an utterance.

These patterns have been used in implementations of embodied conversational agents and robots (see for instance, J. Cassell (1999) or Heylen et al. (to appear)).

This survey of functions and determinants of head movements (still incomplete), shows the variety of factors that are involved. Head movements convey propositional information, they play a role in managing the interaction, are tightly connected with the prosody of speech and they express interpersonal attitudes as well. How can we integrate all these elements into a view on interaction and the design of embodied conversational agents? For this, we have to integrate a linguistic perspective that deals with the syntax and semantics of utterances as well as the organisation of conversations and with a social and psychological perspective.

### 3 Language as Social Action

The number of functions of head movements is bewildering at first side. One way to get a better sense of the function of these behaviours is to consider from a more abstract point of view the nature of conversation or language use and the underlying principles that govern these actions. Particularly, when we take the view that language is a form of social, interpersonal action, one can come to a deeper understanding of the many aspects involved in such simple behaviour as

head movements. Such a view is articulated in Holtgraves (2002) and Clark (1996), for instance. In this section we will summarise Clark's, eclectic, study of language use and hook this up with the various functions of head movements listed above. The major premises of Clark's view that are important to us are the following.

- Language fundamentally is used for social purposes.
- Language use is a species of joint action.
- Language use always involves speaker's meaning and addressee's understanding.
- People need closure on all their actions: lead to people try to ground what they do together.
- Grounding should occur at all levels of communication.
- Many actions come in hierarchies (people do things by doing other things).
- The study of language use is both a cognitive and a social science.

To explain the head movement behaviours and their functions we rely on a couple of key concepts from Clark's perspective on language. Besides the ideas above, this includes the idea of tracks and layers<sup>2</sup>.

**Joint Actions** An important thing to keep in mind when considering behaviours of participants in conversation, is that they are participatory actions that are part of a joint activity carried out by the participants together. In order for such an action to succeed the participatory actions must be coordinated.

"What makes an action a joint one, ultimately, is the coordination of individual actions by two or more people. There is coordination of both *content*, what the participants intend to do, and *process*, the physical and mental systems they recruit in carrying out those intentions." (Clark, 1996, p. 59).

Coordination of actions requires synchronisation of actions. It also requires that each participant closely monitors the actions of the other. And that participants provide feedback of understanding. It is clear from the above, that head movements play an important part in signalling such aspects of the joint

<sup>2</sup>We will not go into the discussion of layers here. Layering is involved in 'pretense' talk: joking, theatrical performance, speaking on behalf of someone else. With respect to head movements, this can be related to McClave's narrative functions.

activity on various levels: from signalling addressee-hood and attention, to interest and even agreement. They are central to the grounding process.

**Action Ladders** The diversity in determinants of head movements is not surprising given that a lot of things happen in conversations at the same time. People do things by doing other things. Coordination works at all these levels simultaneously. Clark distinguishes 4 levels. A communicative act consists of a person A performing some physical action that counts as a signal for something else.

A is executing behavior t for B

A is presenting signal s to B

A is signalling that p for B

A is proposing joint project w to B

Because communicative actions are "joint actions" they are mirrored by actions of the participant B.

B is attending to behavior t from A

B is identifying signal s from A

B is recognizing that p from A

B is considering A's proposal of w

Head movements from listener's provide feedback on all these levels. A listener orients his head to the speaker to obtain more information from facial expression but thereby he is also signalling attention, perhaps understanding and beyond: signalling agreement to the proposal (or joint project) put forward by the speaker.

**Tracks** Clark distinguishes two lines of talk in conversations. The primary track is concerned with "official business", i.e. what the conversation is about. The second track concerns talk (or elements of talk) in the background: talk about the communication itself. Moreover, Clark, remarks that these tracks are orthogonal to the distinction in levels. "The communicative acts in track 2 are used for managing conversation at all four levels of action. When people nod, smile, or say 'uh huh' during another's utterance, they are saying 'I understand you so far,' a signal in track 2 to help achieve closure at level 3." Clark (1996, p. 390). It is immediately obvious from the list of functions of head movements above, that many pertain to this second track. However, not all of them do. The propositional functions, for instance, are mainly used in the official business track.

**Signs and Signals** When listing the 'functions' of head movements a variety of words have been used to characterise the nature of the function: *signal, enhance, anticipate, accompany, express, control, communicate, indicate*.

1. Signal yes or no
2. Signal interest
3. Signal impatience
4. Enhance communicative attention
5. Anticipate an attempt to capture the floor
6. Signal the intention to continue
7. Accompany the rhythmic aspects of speech
8. Express inclusivity
9. Express intensification
10. Express uncertain statements, lexical repairs
11. Control and organize interaction
12. Listing or presenting alternative
13. Communicate the degree of understanding/agreement or support by synchrony of movements
14. Indicate interest, empathy

Hadar et al. (o.c.) consider the question of how the listener's movements signify. Movements that anticipate an action by the listener typically function as cues and signals. On the one hand, they resemble the general movement pattern at the initiation of speech and as such they can anticipate a turn claim precisely in being part of the initiation of speech. But also they often urge the termination of the other's speech. In this sense, they act as a signal for the other: 'I (=listener) want you to stop talking'. "Yes/No" movements are said to operate as symbolic, conventional signals. When one takes a closer look at the various functions of head movements one can also categorise them with respect to the way they mean: whether they are cues, signals, symbols, icons, indices (deictic use).

**Social Actions** Communicative actions are designed to get the audience to do things on the basis of their understanding of what we mean. Illocutionary acts have their origins in social practices. As Argyle points out "each person in an encounter is trying to manipulate the other person, in order to attain his own goals" but on the other hand, we have to take the goals of the other person in mind as well. Holtgraves puts it as follows.

"Not only is language use an action, it is simultaneously an interpersonal action. By interpersonal

action I mean that what we do with language - the actions that we perform (e.g. a request) - have implications for the thoughts and feelings of the involved parties, as well as the relationship that exists between them. Our words are typically addressed to other people, and people are not abstract entities devoid of feelings, goals, thoughts, and values. People's language use - how they perform actions with language - must be sensitive to these concerns. We cannot always say exactly what we mean because we generally do not want to threaten or impose on or criticize our interlocutors." (Holtgraves, 2002, p. 6)

Language is not just a form of joint action designed for the neutral exchange of information. Even exchanging information involves an attempt to change what the other believes. Giving up what one thinks is right, agreeing with what someone else is saying is not a neutral act from a social psychological point of view. In most conversations there is even more at stake for the interlocutors. People use conversations to argue, to negotiate deals or as a prelude to getting more intimate. It should therefore not come as a surprise that people may be offended when an interlocutor apparently does not pay attention by turning his or her head away.

## Conclusion

When one turns to the literature on head movements in conversation, one is faced with a bewildering list of functions and determinants of all the kinds of head gestures that people display during conversations. To get a grasp on the protocols that determine how people move their heads in face-to-face interactions, it is useful to take a step back and consider in more depth what conversations are all about. The basic principles that govern conversation as a joint activity and form of social action can explain most if not all of the patterns of head gestures one may observe.

## References

- Michael Argyle. *Social Interaction*. Atherton press, New York, 1969.
- Michael Argyle and Mark Cook. *Gaze and mutual gaze*. Cambridge University Press, Cambridge, London, New York, Melbourne, 1976.
- Frank J. Bernieri and Robert Rosenthal. *Interpersonal coordination: Behavior matching and interactional synchrony*, chapter 11, pages 401–431. Cambridge University Press, Cambridge, 1991.

- T. Bickmore. *Relational Agents: Effecting Change through human-computer relationships*. PhD thesis, MIT Medialab, 2003.
- T. Bickmore and J. Cassell. *Social Dialogue with Embodied Conversational Agents*. Kluwer Academic, New York, 2005.
- R. Birdwhistle. *Kinesics and Context - Essays on Body Motion Communication*. University of Pennsylvania Press, Philadelphia, 1970.
- Herbert Clark. *Using Language*. Cambridge University Press, Cambridge, 1996.
- D. DiCarlo, M. Stone, C. Revilla, and J.J. Venditti. Specifying and animating facial signals for discourse in embodied conversational agents. *Computer Animation and Virtual Worlds*, 15(1):27–38, 2004.
- A.T. Dittmann. *The body movement-speech rhythm relationship as a cue to speech encoding*, pages 135–151. Pergamon, New York, 1972.
- S. Duncan. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23:283–292, 1972.
- S. Duncan and D. W. Fiske. *Face-to-face Interaction*. Lawrence Erlbaum Associates, Hillsdale New Jersey, 1977.
- N. Freedman. *The analysis of movement behavior during the clinical interview*, pages 153–175. Pergamon, New York, 1972.
- C. Goodwin. *Conversational Organization*. Academic Press, New York, 1981.
- M.H. Goodwin and C. Goodwin. Gesture and coparticipation in the activity of searching for a word. *Semiotica*, 62:51–72, 1986.
- H.P. Graf, E. Cosatto, V. Strom, and F.J. Hang. Visual prosody: facial movements accompanying speech. In *IEEE Conference on automatic face and gesture recognition*, pages 381–386, 2002.
- U. Hadar, T. J. Steiner, and F. Clifford Rose. Head movement during listening turns in conversation. *Journal of Nonverbal Behavior*, 9(4):214–228, 1985.
- U. Hadar, T.J. Steiner, and F. Clifford Rose E.C. Grant. Kinematics of head movements accompanying speech during conversation. *Human Movement Science*, 2:35–46, 1983a.
- U. Hadar, T.J. Steiner, E.C. Grant, and F. Clifford Rose. Kinematics of head movements accompanying speech during conversation. *Human Movement Science*, 2:35–46, 1983b.
- J. Heritage. *Current developments in conversation analysis*, pages 21–47. Multilingual Matters Ltd, Philadelphia, 1989.
- D. Heylen, I. van Es, B. van Dijk, and A. Nijholt. *Experimenting with the gaze of a conversational agent*. Kluwer Academic Publishers, to appear.
- D. Heylen, M. Vissers, H.J. op den Akker, and A. Nijholt. Affective feedback in a tutoring system for procedural tasks. In *ISCA Workshop on Affective Dialogue Systems*, Lecture Notes in Computer Science 3068, pages 244–253, Berlin, 2004. Springer-Verlag.
- T. M. Holtgraves. *Language as Social Action*. Lawrence Erlbaum Associates, Mahwah New Jersey - London, 2002.
- Y. Iwano, S. Kageyama, E. Morikawa, S. Nakazato, and K. Shirai. Analysis of head movements and its role in spoken dialogue. In *ICSLP 96*, 1996.
- S. Prevost J. Cassell, O. Torres. Turn taking vs. discourse structure, 1999.
- W.L. Johnson, P. Rizzo, W. Bosma, S.Kole, M. Ghijsen, and H. van Welbergen. Generating socially appropriate tutorial dialog. In *Affective Dialogue Systems*, pages 254–264, 2004.
- A. Kendon. *Some relationships between body motion and speech: an analysis of an example*. Pergamon, New York, 1972.
- E. Z. McClave. Linguistic functions of head movements in the context of speech. *Journal of Pragmatics*, 32:855–878, 2000.
- Anton Nijholt and Dirk Heylen. Multimodal communication in inhabited virtual environments. *International Journal of Speech Technology*, 5:343–353, 2002.
- C. Sidner, C. Lee, C. Kidd, and N. Lesh. Explorations in engagement for humans and robots. In *IEEE RAS/RSJ International Conference on Humanoid Robots*, 2004.
- K. Smid, I. S. Pandzic, and V. Radman. Autonomous speaker agent. In *CASA 2004*, pages 259–266, MIRALab, University of Geneva, 2004.