Appropriate feedback in asymmetric interactions

Britta Wrede[a,c], Stefan Kopp[b,c], Katharina Rohlfing[b,c], Manja Lohse[a,c], Claudia Muhl[a,b]

[a] Applied Computer Science Group, Bielefeld University, PO Box 100131, 33501 Bielefeld, Germany

[b] CITec (Center of Excellence, Cognitive Interaction Technology), Bielefeld University, PO Box 100131, 33501 Bielefeld, Germany

[c] CoR-Lab (Research Institute for Cognition and Robotics), Bielefeld University, PO Box 100131, 33501 Bielefeld, Germany

Corresponding Author
Britta Wrede
Email: bwrede@techfak.uni-bielefeld.de
Tel: ++49 521 106 2953
Fax: ++49 521 2992
Faculty of Technology
PO Box 10 01 31
33501 Bielefeld
Germany

Abstract

Based on the observation that human-robot interaction is often laborious because a robot's interactional abilities fail to meet the user's expectations, we argue that feedback can play a central role in regulating expectations and mitigating unnecessary disruptions in the flow of conversation. For feedback to be appropriate in this sense, it needs to take situational information into account. This idea stems from interviews with hearing-impaired persons who display perceptual limitations similar to a robot. The results of these interviews indicated that, depending on the goals of the situation, people with hearing impairments used either mediation (clarification) or concealment strategies to keep the interaction going. With this idea in mind, we analyzed human-robot interactions in two different situations—more task-oriented interactions versus more socially driven interactions—and we observed different feedback behaviors in users and their reactions to the robot's behavior. We use these results to derive a scaffold for modeling appropriate feedback in asymmetric interactions (i.e., in human-robot interactions), and briefly discuss consequences for both the design of human-robot interaction and for theories of grounding.

Keywords

Feedback modeling; Grounding; Human-robot interaction

Abbreviations

HRI – Human Robot Interaction

Vitae

Britta Wrede was awarded a Master's degree in Computational Linguistics and her PhD (Dr.-Ing.) in Computer Science from Bielefeld University in 1999 and 2002 respectively. After her PhD, she received a DAAD Postdoc fellowship in the speech group of the International Computer Science Institute (ICSI) in Berkeley, CA. Since 2007 she is heading the research group "Hybrid Society" of the CoR-Lab, Bielefeld University. She is interested in automatic speech recognition and understanding, dialogue modeling for human-robot interaction, and the effects of emotional factors in HRI. She is also working on modeling learning in robots by examining parent-infant interactions.

Stefan Kopp heads the Sociable Agents Group in the Center of Excellence "Cognitive Interaction Technology" at Bielefeld University. He studied Computer Science and Physics at Bielefeld, obtaining his PhD in 2003 for making the virtual human "Max" multimodal. Postdoctoral work at Northwestern University and at the ZiF in Bielefeld investigated cognitive models of communication in embodied systems for intelligent human-computer interaction. Since joining the Artificial Intelligence Group at Bielefeld in 2004, he has continued this research in SFB 673 "Alignment in Communication," the Research Institute for Cognition and Robotics, and the Center of Excellence.

Katharina J. Rohlfing received her Master's degree in Linguistics, Philosophy, and Media Studies from the University of Paderborn in 1997. In 1999 she joined the Graduate Program Task-Oriented Communication at Bielefeld University and obtained her PhD in Linguistics in 2002. She did postdoctoral work at San Diego State University, the University of Chicago, and Northwestern University. In 2006, she became a Dilthey Fellow (Volkswagen Foundation). Since 2008, she is head of the Emergentist Semantics Group, Center of Excellence Cognitive Interaction Technology at the Bielefeld University, and interested in the environmental cues supporting the development of meaning.

Manja Lohse received her diploma in Applied Media Science (*Angewandte Medienwissenschaft*) from the Technical University of Ilmenau, Thuringia, Germany in 2006. After writing her diploma thesis on the "Conception of User-Friendly Personal Service Robots," she joined the research group for Applied Computer Science (*Angewandte Informatik*) in Bielefeld as a PhD student. She is interested in the social aspects of HRI (Human-Robot Interaction) evaluation methods, and users' expectations toward robots.

Claudia Muhl received her Diploma in Sociology from Bielefeld University, Germany in 2005. Her diploma thesis was in the domain of science studies. In March 2005, she joined the Applied Computer Science Group as a PhD student. Her current

research focuses on the social dimensions of human-machine interaction. She is addressing the problem of human interaction with communication robots. Currently, she is working on sociological approaches for analyzing social robots and investigating human-robot communication under the condition of its situatedness and the appearance of unanticipated incidents.

Acknowledgments

This work was supported by the DFG within the SFB 673 “Alignment in Communication” (Projects C1 and C2), by the Volkswagen-Stiftung through a Dilthey-Fellowship and by a grant of Bielefeld University and the Sozialwerk Bielefelder Freimaurer 2005-2007.

1 Introduction

Human-robot interaction (HRI) aims to develop robots that can communicate smoothly with the user. Yet, despite the ongoing technological advances and the increasing capabilities of state-of-the-art robots such as running, juggling, or trumpet playing, an intuitive conversation with a robot still seems to be a very distant prospect. Indeed, interacting with today's robots can often be laborious, because their systems all too often behave in inept and unexpected ways. One major problem—and also one of the greatest challenges when developing conversational robots—is that current systems often do not seem to survey, understand, or even attend to parts of the conversation or the situation in which human-robot interactions occur.

Although a robot's perceptive shortcomings are certainly a hindrance, they should not necessarily prevent effective interaction. As long as users understand the system's capabilities, they can adapt to and compensate for many of its conversational idiosyncrasies. One frequent obstacle to this, however, is the large discrepancy between people's expectations and the robot's actual interactive abilities—something that may differ significantly across situational conditions. Often, after realizing that a robot does not live up to their initial expectations, users do not know any adequate ways to downgrade their mental models of the artificial communication partner. This observation led us to investigate how a robot can signal its limited capabilities and conversational idiosyncrasies, either explicitly through specific communicative behaviors or implicitly through appearance or functioning. Such signaling needs to be adapted to the robot's capabilities in the given situation in the sense that it informs the user precisely about which adjustments should be made. In this paper, we address the question of how a robot can support these adjustments through appropriate feedback.

However, feedback does not just provide information on a communication partner's capabilities. It also serves other functions. The most important one discussed in the literature is the grounding function (Clark and Schaefer, 1989; Clark, 1992) in which the interlocutor's utterances are acknowledged explicitly as heard and understood and possibly even agreed upon. However, in asymmetrical communication situations, this function may be more implicit. This article presents models of feedback and contrasts these with feedback behavior in asymmetric interactions, namely, feedback from artificial agents and feedback strategies in interactions with impaired persons. We then use this as a background to describe observations from two different types of human-robot interaction developed in our labs. We finally use the results of these studies to derive strategies in asymmetric interactions that may be instrumental for the design of communicative agents or robots.

2 Feedback: The background

Originally, the term "feedback" stems from Wiener's (1948) cybernetic notion and describes processes by which a control unit gets information about the effects and consequences of its actions. Because feedback words are often produced during the speaker's contribution, Yngve (1970) has introduced the term "back-channel" to emphasize this permanent bidirectionality of human communication. Other terms and definitions proposed for different nuances of feedback include "listener responses," "acknowledgers," "response words," "conversational grunts," or "roger function" (see, e.g., Allwood et al., 1993; Ward et al., 2000; Allwood et al. in press). A comparative classification is difficult because feedback has been used to define numerous types of behavior used in different contexts, and it can potentially fulfill a variety of functions. As a result, analyzing its semantic/pragmatic or other functions is fairly complex and involves several different dimensions: A "yes" can mean agreement as well as indifference; a "no" may signal surprised agreement at one time and disagreement at another. This multitude of functions expands even further when nonverbal behavior is also considered. A smile can demonstrate sympathy, express self-confidence, or ironically, signal contempt. Likewise, Ehlich (1986) showed that the simple German interjection "hm" carries nine different meanings in dialogue.

With regard to its communicative functions, Allwood et al. (1993: 1) define feedback as "linguistic mechanisms, which enable the participants in a communication to unobtrusively exchange information about four basic communication functions: contact, perception, understanding and attitudinal reactions." In particular, feedback consists of unobtrusive (usually short) expressions that a recipient of information uses to inform the contributor about his or her ability and willingness to communicate (have contact), to perceive information, and to understand it (cf. Table 1 for a list of examples). That is, feedback serves as an early warning system indicating how well speech perception or understanding is succeeding. A feedback utterance at the right time can communicate to the speaker that she or he should, for instance, either repeat the previous utterance and speak more clearly, or use words that are easier to understand. According to Allwood's model, feedback is hierarchically structured on different levels; feedback given at the highest level (attitudinal) draws upon the processing of all lower levels, in particular, contact (perceiving the interaction partner's attention), perception (perceiving the utterance), understanding (processing at a semantic level), and attitude (applying an internal model for evaluation). In contrast, feedback given at the contact or perception level does not necessarily involve understanding or attitudinal processes.

A further potential function of feedback is to communicate whether the recipient accepts the main intention of the contribution (i.e., whether a statement can be believed, a question answered, or a request complied with). Particular emphasis is placed on the mechanisms that underlie the evolution of a successful dialogue and, in particular, the construal of common ground between interlocutors. Following Clark (1992), dialogue should be seen as a collaboration between interlocutors who are cooperating to establish common mutual beliefs. According to this perspective, the driving motivation behind an interaction is mutual understanding, and this can be achieved through stepwise grounding. It involves explicit contributions that form an acceptance phase, signaling that the hearer believes she has understood the content of the previous contribution (Beun et al., 2004).

Function of feedback	**Definition**
Contact	Willingness and ability to continue interaction
Perception	Willingness and ability to perceive expression and message
Understanding	Willingness and ability to understand expression and message
Attitudinal reaction	Willingness and ability to perform other attitudinal reactions to expression, message, or interlocutor

Table 1: Main communicative functions of linguistic feedback according to Allwood et al. (1993).

Thus, in approaches to feedback developed for cooperative human dialogue, the main function is to ground utterances with respect to perception, understanding, and possibly agreement.

2.1 Artificial feedback generation approaches

Almost every existing conversational system that serves as an embodied agent, be it virtual (simulated in computer graphics) or physical (realized in a real robot), models aspects of communicative feedback either implicitly or explicitly. Some approaches explicitly model feedback on a semantic analysis of the interaction. For example, Beun and van Eijk (2004) propose a model that generates elementary feedback sequences at the knowledge level of dialogue participants. Based on an explicit model of the mental states of the dialogue partners, dialogue rules enable their computer system to generate corrective feedback sequences when a user and a computer system differ in their conceptualizations of a particular discourse domain. Affective feedback has been modeled in order to express the agent's own emotional state and how it changes over the course of dialogue (Becker et al., 2004). Alternatively, it can be used intentionally to convey affective states like commiseration. In the Greta character (Poggi et al., 2005), thematic and rhematic parts of a communicative act are assigned an affective state, yielding performative facial expressions that are drawn from large lexicons of codified behavior. AutoTutor (Graesser et al., 2004) is another example of a so-called pedagogical agent that deliberately employs positive ("Great!"), neutral ("Umm"), or negative feedback ("Wrong") to enhance student learning. However, only a few approaches integrate situational information. Heylen et al. (2004), for example, utilize affective feedback to support learning with the tutoring system INES, taking richer information into account such as elements of the student's character, the harmfulness of errors made, and the emotional effects of errors.

Earlier systems have already acknowledged feedback as an integral part of communicative behavior, stressing its importance in, for instance, managing the flow of conversation. Most of these approaches work on the contact or perceptual level. The Gandalf system (Thórisson, 1996) employs pause duration models to generate agent feedback. For example, it gives verbal back-channel utterances or head nods after a silent pause of a certain duration (110 ms) in the speaker's utterance. Gandalf simulates turn-taking behavior by looking away from the listener while speaking, returning his gaze when finishing the turn. The REA system (Cassell et al., 1999a b) also uses a pause duration model and employs different modalities for feedback (head nods, short feedback utterances). Like Gandalf, Rea looks away at the beginning of her turn and returns her gaze to the listener when a change in speaker role is intended.

BodyChat (Vilhjálmsson et al., 1998) is a system that demonstrates the automation of communicative behaviors in avatars for users who communicate via text. Their avatars automatically animate attention, salutation, turn-taking, back-channel feedback, and facial expression, and they also execute simple body functions such as eye blinking. Feedback behavior selection was boiled down to rules such as "RequestFeedback by Looking or RaiseEyebrows" or "GiveFeedback by Looking and HeadNod."

In recent years, several systems have been developed to improve predictions on the right timing for feedback. Ward and Tsukahara (2000) describe a pause-duration model that also incorporates prosodic cues based on the best fit to a speech corpus. Takeuchi et al. (2004) augment this approach with incrementally obtained information about word classes. In addition to analyzing prosody information to extract proper feedback timing, Fujie et al. (2004) employ a network of finite state machines including one that maps recognized words onto content for possible feedback before the end of the utterance. Their model is implemented in the conversational robot ROBISUKE that also uses short head nods for feedback.

Evaluation studies have tested the effects of modeled feedback behaviors for many years. Experiments with the Gandalf agent revealed that the presentation of content-related feedback (successful question answering or command execution) together with so-called envelope feedback (such as gaze and head movement for turn-taking or co-verbal beat gestures) led to smoother interactions with fewer user repetitions and hesitations. Additionally, although the language capability of the system was identical to that in other conditions, it was rated as being better (Cassell et al., 1999a). However, other evaluation studies showed that the ability of the models in common use to predict feedback is only limited (e.g., Cathcart et al., 2003).

Gratch and colleagues (2006) describe a recent experiment on multimodal, yet nonverbal agent feedback and its effects on the establishment of rapport. Their "Rapport Agent" was designed to elicit rapport while listening and giving feedback to a human who is describing a previously watched cartoon sequence. This system analyzes the speaker's head movements and body posture with a camera and implements Ward et al.'s (2000) pitch cue algorithm to determine the right moment for giving feedback by head nods, head shakes, head rolls, and gaze. In comparison with random head moves and posture shifts, the system seems to elicit an increased feeling of rapport in human dialogue partners: They addressed significantly more words and told longer recaps to the Rapport Agent. Moreover, their self-report evaluation showed higher ratings of the agent's understanding of the story and a stronger feeling of having made use of the agent's feedback.

Table 2 summarizes these approaches in terms of the feedback levels defined by Allwood et al. (1993). It can be seen that most of the technical approaches to feedback modeling reviewed above address the levels of contact and perception.

Main communi-cative function	Strategy	Function	Example	Authors
Attitudinal	Show affective state	Convey information about own state	Smile	(Becker et al., 2004),
Attitudinal	Show affective state	Convey information about own state	"Well done"	(Poggi et al., 2005)
Attitudinal	Evaluate student's performance	Convey information about other's performance	"Great" "Uhmm" "Wrong"	(Heylen et al., 2004)
Understanding	Communication of mental states	Convey information about own internal state	Corrective feedback sequences	(Beun & van Eijk, 2004)
Perception / Contact	Turn-taking	Keep conversation	Head nod after long	(Thórisson, 1996)

		going / flow	pause	
Perception / Contact	Turn-taking	Keep conversation going / flow	Backchannel utterance after long pause	(Cassell et al., 1999a), (Cassell et al., 1999b)
Perception / Contact	Attention	Being attentive to talking user	Raise Eyebrow	(Vilhjálmsson et al., 1998)
		Being attentive to object	Gaze at object	
Perception / Contact	Turn-taking	Request Feedback	Looking, raising eyebrows	(Vilhjálmsson et al., 1998)
		Give Feedback	Looking, head nod	
Perception / Contact	Rapport Building	Give Feedback	Head nods, shakes, and gaze	(Gratch et al., 2006)
Perception / Contact	Eye contact	Hold the floor	Looking away	(Thórisson, 1996), (Cassell et al., 1999a), (Cassell et al., 1999b)
		Give the floor	Returning gaze	
Perception / Contact	Turn-taking	Continuation	Provide feedback based on pause duration model	(Ward & Tsukahara, 2000), (Takeuchi et al., 2004)
Perception / Contact	Turn-taking	Continuation	Provide feedback before end of utterance	(Fujie et al., 2004)

Table 2: Feedback strategies, their functions, and feedback level applied in artificial communicative agents.

However, it should be noted that some approaches explicitly take this hierarchy of feedback levels into account (Brennan et al., 1995; Larsson, 2003). In Brennan et al. (1995), the levels on which feedback is given are selected according to how well the interaction is proceeding. If there are many clarifications, the system provides feedback at every level; that is, it signals explicitly that it has heard the user's utterance, has parsed the words, has understood, and will take action. This enables the user to understand where the system may have made an error. In Larsson (2003), these feedback levels are modeled implicitly. However, in both cases, the user's utterance is subject to a hierarchy of steps of analysis, and failure to understand is signaled immediately to the user, providing feedback on different levels.

As this summary indicates, many artificial feedback generation approaches focus on context-independent generation rules that disregard situational information; that is, they produce feedback behavior mechanistically in the form of a certain reaction in response to a limited number of user cues. These feedback models operate mainly on a perceptual or contact level. Yet, as evaluations have shown, they are able to exert a positive influence on the user's perception of the system and even the communicational flow. On the other hand, these models show a risk of not being perceived as natural. One example of this is the prediction of timing that does not always match human timing satisfactorily (e.g., Cathcart et al., 2003). In contrast, few artificial approaches try to analyze the user's utterance on a semantic level in order to provide a grounding-oriented feedback. This makes it necessary to ask what a system might gain by employing deep semantic analysis when shallow feedback already yields positive user evaluations. Is it possible to distinguish situations in which shallow processing suffices from situations that need understanding at a deeper semantic level?

Moreover, the models of communication outlined above commonly assume that human communication is effective and successful. This might be the case in most everyday situations between fully competent interlocutors. However, when we face an unexpected and atypical partner such as a person with a disability or an artificial agent, communication will develop differently. For many people, such an interaction is associated with discomfort. The unfamiliar and atypical appearance of people with disabilities or their physical limitations and idiosyncrasies in communication behavior trigger anxiety and defense reactions. This 'contact problem' is relevant for feedback design, because artificial systems can be seen as atypical partners: Their communicative capabilities have to be evaluated by interaction partners in order to elicit adaptive processes (Vollmer et al., submitted). For most people, a robot is an atypical partner; in first encounters, it is not clear which perceptual sensors and communicative abilities the robot brings into the interaction. Some of these capabilities may be inferred in part from the robot's appearance (e.g., its anthropomorphic form; see Lohse et al., 2008b). Therefore, it is important to design the first contact and to make the robot's abilities as transparent as possible to increase the user's confidence in the interaction (Lohse et al., 2008a). In the following, we shall define not only those situations in which more transparency in the robot's behavior is necessary but also those in which less transparency may be beneficial on the basis of observations of asymmetric human interactions.

2.2 Feedback strategies in interactions with impaired persons

Although most studies conducted from the perspective of interaction analysis investigate conversations between typical interaction partners, some studies, such as Goodwin (2003), examine atypical interlocutors. In a 'first contact' situation with an atypical partner, behavioral strategies from typical communication are of little use, and this may elicit feelings of anxiety. These feelings become compensated during the ongoing interaction as participants adapt to each other and generate interaction strategies online. Constructing and negotiating mutual understanding ("co-construction," Collins and Marková, 1999) seems to be the goal of a successful communication. Collins and Marková (1999) argue that in atypical situations, it is mostly the nondisabled rather than the disabled person who tends to adopt a leading role in guiding the conversation and its pragmatic aspects (e.g., determining topic and managing turn-taking). This is the way in which collaborative knowledge is established (Brewer et al., 1991). This leads us to ask how grounding can actually be achieved in asymmetric interactions in which one participant may not be able to follow the processes necessary for grounding.

In order to derive strategies for a robot to produce appropriate feedback behavior, we adopted the view that a robot can be considered as an untypical communication partner. Therefore, we looked for strategies of communication maintenance and expectation regulation in interactions between physically healthy and impaired humans—what we call asymmetric interactions—by conducting a group discussion with people who used to work professionally with disabled people (with hearing, speech, and mental impairments) and who are now physically impaired themselves.

Six deaconesses who had worked as nursing staff for old people prior to their retirement, and three researchers were present at the discussion. While one researcher managed the audio recordings, two participated actively in the discussion by asking questions and summarizing contributions when this seemed appropriate. The discussion was planned carefully in advance and guided by some set questions on how one notices that an interlocutor has limitations in communication, how one adapts to these limitations, how one communicates one's own impairments, and how one expects the other persons to adapt. We informed participants that the purpose of our investigations was to design appropriate feedback in an asymmetric communication in which two differently competent speakers face each other.

Audio recordings and notes from this discussion were analyzed to detect strategies that the impaired persons and the nursing staff reported as having a facilitative function in the conversation. We also analyzed the discussion for strategies that can be used to recognize and compensate limitations in the interaction partners.

2.2.1 Feedback strategies of impaired persons

The analysis of the interview indicated that impaired persons actually do have a set of strategies to ensure the content or to maintain the process of an interaction. Speech- or hearing-impaired persons tend to apply two strategies to keep a communication ongoing: mediation and concealment of their communication impairment. An example of a mediation strategy is (a) to make difficulties in comprehension explicit and to offer specific solutions, for example, by saying "Sorry, I can't hear you well, could you speak louder / slower / in my direction?" Interestingly, these strategies tend to focus on the perception level, implying that the impairment does not affect an individual's general ability to understand. Examples of concealment strategies are (b) to pretend comprehension nonverbally by, for example, nodding even though comprehension is impaired, or (c) to utter verbally meaningful, commonplace phrases such as "How interesting!" even when one has barely understood what the partner has said.

2.2.2 Feedback strategies of nursing staff

Nursing staff apply the following strategies to maintain communication: to sensitively recognize the limitations and difficulties within an interaction and to compensate them. The most important indicators of communicative limitations in an interaction partner were reported to be (d) facial expression, such as a visible effort in trying to see or hear, (e) eye-gaze and head-turns that can signal an effort, or ongoing comprehension by nodding; and (f) no reaction when feedback was expected. Even though the strategies (d) to (f) were mentioned as options when monitoring ongoing communication in the discussion, the deaconesses emphasized experience as the basis for establishing and maintaining contact with an impaired person. Hence, it is almost impossible to identify impairment from initial feedback. It is necessary to approach and probe the communicative potentials and limitations during the conversation.

When applied to feedback design rules for artificial interaction partners, this would indicate that users need a certain training time in order to adapt to the idiosyncrasies of the system and to learn to distinguish between 'true' feedback and concealment. However, this might not be as difficult as it appears, because the nursing staff reported subtle but consistent cues (see d to f above). On the other hand, detecting an attempted concealment does not necessarily lead to an attempted clarification (e.g., a suggestion to repeat). According to Tannen (1984), clarification attempts may be face-threatening and should therefore be used with care. Context information, such as the history of the interaction or its task orientation, should be taken into account when a concealment attempt has been detected.

Main Communicative Function	**Strategy**	**Function**	**Example**
Understanding	Search for signs of nonunderstanding	Detect concealment	Observation of facial expression, eye gaze / head turns, missed reactions
Perception	Offer specific solutions in terms of behavior suggestions	Clarification / Mediation	"Speak up!" "Look in my direction!" Nursing staff:

			"Should I repeat?"
Perception	Nonverbal feedback for understanding even when not understood	Continuation / Concealment	Head nodding "Mhm"
Perception	Commonplaces	Continuation / Concealment	"How interesting" "That's good to hear"

Table 3: Feedback strategies and their functions as reported for interactions with (sensorily) impaired persons.

2.2.3 Conclusion

Apparently, participants in an asymmetric interaction use different strategies when confronted with understanding problems. On the one hand, they employ mediation strategies that serve a similar function to the clarifications in Clark's notion of grounding in order to achieve mutual understanding, and they employ these preferably in situations in which a task has to be solved (e.g., collaborative discourse or task tutoring). On the other hand, concealment strategies are applied frequently that contradict current ideas on communication in which the goal is to solve a joint task. We hypothesize that concealment helps impaired persons to stay engaged in an interaction by maintaining contact (i.e., attention, reciprocal interaction), even if they do not fully comprehend. Such concealment strategies are more likely to remain undetected or at least undisclosed in cases in which no common task has to be achieved. Indeed, it seems that the use of concealment strategies depends on the history of the interaction and the situational goals .

In contrast, people who are used to interacting with impaired persons possess strategies for detecting concealment attempts, and they offer implicit or explicit clarifications in order to resolve problems of nonunderstanding. Again, the application of these strategies seems to depend on the goals of the current situation.

Based on these results, we define the goals of the situation in terms of two dimensions: to what degree an interaction takes place in order to fulfill social functions, that is, to stay in contact, show mutual interest, and so forth (social orientation); and to what degree the interaction takes place in order to jointly solve a task (task orientation). Hence, the goals of a situation would be easy to control in an experimental setting by specifying different tasks that participants have to perform.

3. Feedback effects in HRI

In the following, we shall present analyzes of HRI taking place under different conditions. First, we shall analyze interactions with an embodied, virtual agent in a socially driven situation in which users were instructed to talk to the agent without being given a specific task. Second, we shall present results on interactions with a robot in a task-oriented interaction in which users were instructed to show the robot around an apartment. Based on these analyzes, we shall report different uses of feedback strategies and formulate hypotheses about the appropriateness of feedback depending on the situation operationalized by the instructions.

Fig. 1

Figure 1. Analyzed scenarios: The two analyzed scenarios in the situation space as defined by the dimensions social orientation and task orientation.

3.1 Socially driven interaction with an embodied virtual agent

As an example of socially driven interaction, we analyzed interactions with the virtual agent MAX developed to investigate the social effects of embodied agents on user perception. Max is a virtual anthropomorphic agent designed to simulate features of human face-to-face communication for interaction with machines (Kopp et al. 2005). Max has been installed in the Heinz Nixdorf MuseumsForum (HNF), a public computer museum in Paderborn (Germany), where he tries to engage visitors in conversations during which he provides them with background information about various topics of interest.

Fig. 2

Figure 2. The virtual character Max: The picture shows the setup in the Heinz-Nixdorf Museum where interaction sequences were recorded. Max produces facial expressions and hand gestures corresponding to its verbal output. Interactions are mainly socially driven.

We analyzed the content of user utterances in one week of log files containing 205 dialogues in order to find out what communication strategies people used in this socially driven interaction. The instruction given to the users was to talk to the agent without a specific task. This resulted in what we call a socially oriented situation, and this was confirmed by the observation that participants tested the system's capabilities by asking difficult questions or giving unrealistic answers, thus demonstrating that they were communicating for the sake of the interaction.

In this scenario, we analyzed the users' feedback strategies in terms of their social function. Although we are focusing on how to design appropriate robot feedback, we chose this method in order to analyze how users react to an artificial system in a situation in which a task solution is not the major issue.

Results indicate that users showed a high degree of proactiveness, with 35.9% of their utterances being rated as proactive. They also demonstrated a strong willingness to interact, with 75.0% of their answers to the agent's questions being expedient and inconspicuous. Thus, users showed very active communicative behavior in order to gain and maintain contact with the agent. Overall, the users' utterances contained more positive than negative feedback. Interestingly, when interacting with the virtual agent, they often referred to human-like concepts that one would not expect to make sense when interacting with a computer or a robot (e.g., "Are you in love?"). They also used a high number of the conventionalized dialogue routines commonly found in small talk such as greetings (57.6% of all encounters) or farewells (29.8% of all encounters). In contrast, although a large number of requests had to be repeated (241), there were hardly any requests for clarification (5).

These results indicate that participants in the socially oriented interaction analyzed here chose topics and utterances that are easy to answer without deep semantic analysis. Moreover, participants' strategies are manifold and borrow extensively from small-talk routines. In contrast, the small amount of feedback with the function of securing understanding indicates that understanding is not the main goal of this type of interaction.

3.2 Task-oriented interaction with a mobile robot

We investigated task-oriented communication by analyzing interactions with the mobile robot BIRON (Bielefeld Robot Companion). This is a multimodal situated robot that can learn the spatial environment as well as the names and visual appearances of objects (Hanheide and Sagerer, 2008). BIRON understands not only spoken utterances but also co-verbal deictic references to objects in the scene. Moreover, the robot can carry out mixed-initiative dialogues. It uses several feedback channels such as speech, movement, the display of an animated robot character, and a moving pan-tilt camera.

Fig. 3

Figure 3. The mobile robot Biron: The user's task is to show the robot the layout of an apartment and several prespecified objects. The robot can see the user through its self-moving pan-tilt camera mounted on the top of the display. It can localize the sound source of a speaking voice through the stereo microphones below the display, observe the gestures through the camera below the microphones, and track the user's legs with the laser range finder (blue box at the bottom of the robot).

The results reported here are taken from experiments embedded in the so-called home tour scenario. This scenario simulates a situation in which a robot is delivered to the home of a user who has no experience in interacting with a robot. By interacting with the user, the robot can acquire knowledge about its new environment such as the rooms and various objects. Participants were given a short introduction to the robot's basic functionalities before being instructed to show it several rooms and objects.

We analyzed the users' task solutions on the system and interaction level of the interaction. The system level includes automatically generated log files of the components such as motor commands, robot speech, and their interplay. However, the present analysis focused on an evaluation of the interaction level. We used the annotation tool ELAN to annotate human speech, movement, gestures, and facial expressions and synchronize them with the video recording of the interaction.

We then conducted a task analysis to identify prototypical interaction scripts for each task (i.e., greeting, showing rooms and objects, guiding, farewell) along with deviations from these scripts. This enabled us to recognize when the interaction went smoothly and when it was disturbed. Moreover, we were able to see how the users tried to maintain the interaction flow. This article presents only some outstanding examples. A complete and detailed analysis can be found in Lohse et al. (2009).

One frequent type of feedback related to the turn-taking level of the interaction: Users often tried to get the robot's attention by, for example, trying to position themselves in front of the robot camera, thus using the camera's feedback-signaling function. This was the case in almost all interactive situations such as before and during a so-called following sequence in which the robot was supposed to follow the user. If positioning did not succeed, many users started waving. Similar user behavior has been reported from interactions with a virtual robot (Muhl et al., 2007), indicating that gaze is indeed a crucial feedback behavior. Observations of the users' reactions to disturbances in the perceived attentional behavior of the robot indicated that the robot's gaze was interpreted as a meaningful cue for its focus of attention. When the robot's attention shifted to an unpredicted point, users tried to regain it (e.g., by waving their hands or moving their body), continued to just look at the robot, or, finally, turned to the experimenter as interaction partner. Although individuals chose different types of behavior, the overall repertoire of reactions across the whole range of participants was very limited. The importance of appropriate gaze behavior has also been demonstrated in virtual agents in which task achievement was affected negatively when the gazing behavior did not meet the users' expectations (Heylen et al., 2002). From these findings, we can conclude that the agent's gaze is an important prerequisite for an interaction, and that users rely on it to ensure that the system is ready to act. Therefore, this should be taken into account when modeling appropriate feedback generation.

As well as examining the responses to the robot's gazing behavior, we also performed a detailed analysis of sequences with disfluencies in the conversation. The term 'disfluency' refers to situations in which the expectations of one interaction partner, as perceived by a trained annotator, are not matched by the robot’s behavior. Because the annotator had previously been trained to analyze interactions with this specific robot, she was experienced in judging the user's expectations when, for example, giving a specific command. Knowledge about the functioning of the robot system was important in order to understand the robot's reaction.

Results indicate that many disfluencies occurred in the context of unclarity about the robot's functionalities. Most importantly, the users requested explicit information about how to communicate with the robot. This addressed, on the one hand, how to handle turn-taking, because waiting times of 20 s or more occurred in cases when the robot had not signaled that it was ready to receive a command. On the other hand, it concerned explicit action (or utterance) proposals by the robot in order to inform the users about what action possibilities were available in a specific situation. This was also the case for possible solutions to technical problems. For example, when the robot got physically stuck in a narrow area, users were very compliant when the robot asked them to give it a push. Thus, although giving explicit information about the robot's turn-taking or feedback expectations does not appear to be an elegant and smooth strategy for interaction, users requested it in order to maintain the flow of interaction, and it did, indeed, prove to increase the interaction flow.

We also observed that asking the user to repeat an utterance did not lead to irritation. More importantly, users were very sensitive to the robot's feedback signaling its understanding of certain actions. For example, in a first iteration of our study, the robot reacted to a deictic reference by the user (e.g., "Look, this is the cupboard") by moving its camera toward the (estimated) object pointed out by the user and answering "That's interesting. I really like it." However, users did not accept this as sufficient proof of understanding, and they often asked questions like "Really?" or "Can you repeat the name of the object?" Users displayed disbelief and irritation by using facial expressions to indicate that they did not think that the robot had understood correctly. Thus, the affective evaluation of the robot ("I really like it") suggested a higher level of understanding than users would attribute to the robot after having experienced a certain history of interaction. This behavior slowed down the interaction considerably and interrupted the flow. Therefore, we replaced this by programming the robot to (a) look longer at the object and (b) repeat the object's name (e.g., "A cupboard, I've seen it"). This behavior was accepted by the users and increased the robot's credibility significantly.

From our studies with the robot BIRON, we can conclude that the following feedback strategies in the robot support a smooth interaction: giving explicit action proposals, suggesting solutions, asking for repetition, and giving convincing proof of understanding (i.e., repeating important information). All these patterns of feedback behavior aim to facilitate the completion of a **joint task**, thus emphasizing the task-directedness of this interaction.

3.3 Conclusion

These results from human-machine interactions in two different situations lead us to the following conclusions: (a) Socially driven interactions reveal more social types of feedback such as dialogue routines or the use of anthropomorphic concepts and attributes, but hardly any clarification of information. Although this is supported by our observations on feedback strategies in interactions with impaired persons, this result contradicts grounding theories based on the assumption that the goal of an interaction is the completion of a joint task (e.g., Clark, 1992). However, these results could be integrated into such notions of grounding by interpreting socially oriented interactions as aiming to maintain contact and keep the conversation flowing. (b) In task-oriented interactions, many disruptions occur because users need more information about the way the system works before they understand its behavior or can carry out clarification sequences. In these interactions, feedback focuses on the joint task and its completion. In this collaborative work, gaze is a crucial feedback behavior. It has dramatic effects on the ongoing interaction because it provides the basis for joint attention and, hence, for task completion.

4 Toward a model of appropriate feedback

4.1 The importance of 'flow'

Combining our observations of human-robot interactions with the human-human interactions involving people with impairments, we find: (a) In asymmetric communication settings, interlocutors can employ feedback to fulfill two main functions: mediation or concealment. (b) Both functions serve a higher ranking purpose, namely, to ensure successful communication, and this crucially seems to involve maintaining a 'flow.'

According to Clark (1992), interactions (at least between nonimpaired human interaction partners) follow a structure of presentations and acceptance that gradually builds a common ground of shared beliefs (compare Fig. 1). Such a structure is important when the interaction is directed toward a certain goal as in task-oriented interactions. Strategies of mediation, as reported above, would be interpreted as clarifications; that is, a repetition of presentations in order to solve an understanding problem before acceptance can be given.

Fig. 4

Figure 4. Communicational flow in task-oriented interaction. When a task has to be accomplished, the grounding primitives acceptance and presentation need to be supplemented by clarifications.

However, concealment strategies are not anticipated in such a model in which the driving motivation is to achieve a common goal. Thus, concealment strategies only make sense in an interaction in which the goal is not to achieve a task (compare Fig. 3). The function of concealment is to maintain the structure of presentation and acceptance, and thus the flow of interaction. Note that concealment is not random, but has to take the dialogue structure into account and also follow turn-taking rules.

Fig. 5

Figure 5. Communicational flow in a socially driven interaction: No clarifications need to be carried out in interactions when no task has to be accomplished. The correct sequence of presentation and acceptance suffices to maintain the communicational flow.

4.2 A model of appropriate feedback generation

Based on our observations of different feedback strategies in both HRI and asymmetric human-human interaction, we can define variables that influence the choice of feedback strategy in such interactions. We have identified concealment versus mediation as feedback strategies in the case of nonunderstanding. Variables influencing the choice between these two strategies relate to the interaction situation. For our interpretation, we have chosen the dimensions of task orientation and social orientation. The following two sections consider these variables in more detail, starting with feedback strategies that are especially suitable for concealment feedback and ending with a rough scaffold to model the decision process for feedback strategy selection.

4.2.1 Feedback strategies for concealment

Feedback strategies for concealment are related to the feedback levels of contact and perception. Feedback given at these levels, such as "mhm" or "yes," is easy to interpret as signs of understanding without any need for explicit lies. This can be seen in the strategies for concealment reported by impaired persons (cf. Table 4): They give nonverbal feedback such as nodding or smiling, or offer commonplace phrases. This reveals a striking similarity with strategies reported for artificial systems (cf. Table 3), suggesting that even though artificial agents work with more or less sophisticated computational models, they are actually using concealment strategies. The types of interaction suggest that these interactions are highly socially driven with a weak focus on the achievement of a joint goal. In contrast, the few feedback strategies pertaining to levels of understanding or attitude have been implemented for interactions in task-oriented situations. Thus, it appears that artificial agents also use concealment strategies—which is reasonable in some ways, given that they also tend to have difficulties in analyzing utterances especially in the unrestricted interactions that are typical for virtual social agents. However, in contrast to impaired humans, artificial agents do not know in which situations it is appropriate to apply concealment strategies and in which it is better to acknowledge nonunderstanding.

4.2.2 When to apply feedback strategies?

Returning to the data from our human-human interaction study, we have seen that impaired persons use both concealment and mediation strategies. To model this behavior, it is necessary to determine the type of situation: whether the ongoing interaction is pursuing a joint goal versus whether it is more socially driven. This could be done easily by simply letting the robot define the situation: For example, it could initiate an interaction by offering services (task orientation) or by initiating small talk (social orientation). More complex perceptual capabilities would be needed to monitor changes in the goals of an ongoing interaction. This means constructing an internal model of the goals, structure, and status of the ongoing discourse in order to decide upon the feedback strategy. Such a model would need to take the history of the interaction into account and avoid, for example, any excessively long sequence of clarifications. Features that may help to determine the degree of social orientation automatically in the ongoing interaction may be based on statistical lexical analyses of the dialogue history by, for example, using look-up dictionaries containing common social dialogue routines. Another feature may be the internal state of the robot and its history incorporating information about which functions have been activated. Future research could examine these conversational capabilities together with the ability to determine the exact conditions of a task-oriented conversation.

Given such a model of the discourse context, feedback can be generated according to the process depicted in Figure 3: In the case of nonunderstanding, the system has to check the current type of interaction in order to determine whether to initiate a clarification sequence or pursue a concealment strategy.

Simply analyzing the type of interaction is not sufficient for a smooth conversation. If a system is to update the internal discourse model and monitor its own performance, it needs to take the interaction partner's reaction into account. If the partner accepts the concealment, the system has to note that the utterance was successful and the interaction can proceed. But if the interaction partner scrutinizes the answer, the robot has to enter into a clarification process and needs to update its discourse model in order to avoid further wrong judgments about the current discourse situation. This monitoring is essential, because the interactions with BIRON have shown that users do not accept every positive answer from the robot. If they think that the system does not really understand, they ask it again to obtain proof of understanding (scrutiny).

Fig. 6

Figure 6. Decision tree for modeling feedback in asymmetric interactions: Rough sketch of decision process to determine the type of feedback in an ongoing interaction. (R: indicates a contribution by the robot; U: indicates a contribution by the user)

5 Conclusion

This article argues that feedback plays an important role in enabling a robot to communicate its capabilities. Our observations from actual human-robot interaction as well as from interviews with people with impaired hearing and other physical and mental impairments indicate that feedback needs to regulate user expectations while mitigating unnecessary disruptions of the flow of conversation. We argue that in order to be able to produce appropriate feedback in an asymmetric interaction, a robot has to take the goals of the interaction into account. In our definition, the goals of a situation unfold along two dimensions: task orientation and social orientation. These goals are particularly important in cases of nonunderstanding, because they allow the robot to decide whether to enter a clarification sequence or to employ a concealment strategy. More task-oriented interactions require clarification, and this has to make use of all available communicative resources in order to regain mutual understanding. Such behavior has been shown in human-robot interactions when the users are very willing to comply with the robot's requests (e.g., to follow very narrow interaction patterns). More generally, users show a very high degree of compliance in order to reach a common goal, for example, in cases when the robot has to ask for help. Thus, in task-oriented interactions, feedback requesting clarification and help is acceptable for users and thus most appropriate.

In contrast, in more socially driven interactions, users do not use clarification strategies themselves and do not expect them from the robot either. Instead, we see a predominant use of strategies of concealment in order to keep the flow of interaction going. However, the robot has to make sure it does not use a concealment strategy in the wrong situation. Users scrutinize a robot's concealing responses more often when the task is to achieve a joint goal. Moreover, concealment needs to make use of all the actuators available to the robot. Verbally signaling understanding will not be convincing if the system does not, for example, display joint attention through its camera movement. This has consequences for the overall system architecture, because the process of grounding (or concealment of problems within it) needs to be tightly integrated at a deep level in the system on which all actuators and sensors can be included in this process.

These conclusions are only partly in line with current theories of dialogue management. For example, grounding theories state that communication is driven by the desire to reach a joint goal. This goal has commonly been interpreted as a task-oriented one. Therefore, grounding has basically been analyzed in task-oriented situations. However, it should be noted that Clark (1992: 259), for example, states that the grounding criterion is based on "current purposes," and this allows for ample interpretation such as a differentiation between a social and a task orientation. As we have shown, the goal of a communication is not necessarily tied only to a task. It may also serve the desire of agents to remain in contact through a reciprocal interaction with functioning turn-taking and a continuous flow of communication. Such more socially driven interactions have been analyzed generally as social dialogue (in terms of phenomena like small talk or conversational frames) and within the area of dialogue management for artificial social agents. However, in contrast to existing feedback generation approaches that also use concealment strategies, we propose the explicit use of concealment strategies in order to behave in accordance with situational user expectations.

The differentiation between socially oriented and task-oriented goals allows us to analyze grounding in different situations. One situation in which concealment (a strategy that apparently contradicts the notion of grounding) may be employed is when temporary problems of perception or understanding arise, but these are assessed as not being too severe or as probably resolvable once more information has been elicited. Alternatively, concealment may simply be the result of a predominant wish of the communication partners to remain engaged in an interaction with expected desirable social outcomes. This reflects the view that not only the achievement of a task but also the wish for social interaction is a driving force behind communication. Whether this hypothesis also holds for human-robot interaction will require further research, and it may help to shed light on the relationship between the driving forces of dialogue.

References

Allwood, Jens, Kopp, Stefan, Grammer, Karl, Ahlsen, Elisabeth, Oberzaucher, Elisabeth, Koppensteiner, Markus, in press. The analysis of embodied communicative feedback in multimodal corpora – a prerequisite for behaviour simulation. Journal of Language Resources and Evaluation.

Allwood, Jens, Nivre, Joakim, Ahlsén, Elisabeth, 1993. On the semantics and pragmatics of linguistic feedback. Journal of Semantics, 9 (1), 1–26.

Becker, Christian, Kopp, Stefan, Wachsmuth, Ipke, 2004. Simulating the Emotion Dynamics of a Multimodal Conversational Agent. In: Proc. Affective Dialogue Systems: Tutorial and Research Workshop. Series: Lecture Notes in Artificial Intelligence 3068. Springer, Berlin/Heidelberg, pp. 154–165.

Beun, Robbert-Jan, van Eijk, Rogier M., 2004. Conceptual Discrepancies and Feedback in Human-Computer Interaction. In: Proc. Dutch directions in HCI. ACM Press, New York, pp. 13-16.

Brennan, S. E., Hulteen, E., 1995. Interaction and feedback in a spoken language system: A theoretical framework. Knowledge-Based Systems, 8, 143–151.

Brewer, John D., McBride, Gill, Yearley, Steven, 1991. Orchestrating an encounter: a note on the talk of mentally handicapped children. Sociology of Health & Illness, 13, 58–67.

Cassell, Justine, Thórisson, Kristinn Runar, 1999a. The Power of a Nod and a Glance: Envelope vs. Emotional Feedback in Animated Conversational Agents. International Journal of Applied Artificial Intelligence, 13 (4–5), 519–538.

Cassell, Justine, Bickmore, Timothy W., Billinghurst, Mark Nathan, Campbell, Lee Winston, Chang, K., Vilhjálmsson, Hannes Hogni, Yan, Hao, 1999b. Embodiment in Conversational Interfaces: Rea. In: Proc. Human Factors in Computing Systems (CHI): 520–527.

Cathcart, Nicola, Carletta, Jean, Klein, Ewan, 2003. A shallow model of backchannel continuers in spoken dialogue. In: Proc. European Chapter of the Association for Computational Linguistics (EACL10): 51–58.

Clark, Herbert H., Schaefer, Edward F., 1989. Contributing to discourse. Cognitive Science, 13, 259–294.

Clark, Herbert H., 1992. Arenas of language use. University of Chicago Press, Chicago.

Collins, Sheila, Marková, Ivana, 1999. Interaction between impaired and unimpaired speakers: Inter-subjectivity and the interplay of culturally shared and situation specific knowledge. British Journal of Social Psychology, 38. 339–368.

Ehlich, Konrad, 1986. Interjektionen (Interjections). Niemeyer, Tübingen, Germany.

Fujie, Shinya, Fukushima, Kenta, Kobayashi, Tetsunori, 2004. A Conversation Robot with Back-channel Feedback Function based on Linguistic and Nonlinguistic Information. In: Proc. Autonomous Robots and Agents (ICARA): 379–384.

Goodwin, Charles, 2003 (Ed.). Conversation and Brain Damage. Oxford University Press, Oxford.

Graesser, Arthur C., Lu, Shulan, Jackson, Gerorge Tanner, Mitchell, Heather Hite, Ventura, Mathew, Olney, Andrew, Louwerse, Max M., 2004. A tutor with dialogue in natural language. Behavioral Research Methods, Instruments, and Computers, 36, 180–193.

Gratch, Jonathan, Okhmatovskaia, Anna, Lamothe, Francoise, Marsella, Stacy, Morales, Mathieu, van der Werf, R. J., Morency, Louis-Philippe, 2006. Virtual Rapport. In: Proc. Intelligent Virtual Agents (IVA), Series: Lecture Notes in Computer Science 4133. Springer, Berlin/Heidelberg, pp. 14–27.

Hanheide, Marc, Sagerer, Gerhard, 2008. Active Memory-based Interaction Strategies for Learning-enabling Behaviors. In: Proc. Robot and Human Interactive Communication (RO-MAN): pp. 101–106.

Heylen Dirk, van Es, Ivo, van Dijk, Betsy, Nijholt, Anton, van Dijk, E. 2002. Experimenting with the gaze of a conversational agent. In: Proc. International CLASS Workshop on Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems, 28-29 June 2002, Copenhagen, Denmark. pp. 93-100.

Heylen, Dirk, Vissers, Maarten, op den Akker, Rieks, Nijholt, Anton, 2004. Affective Feedback in a Tutoring System for Procedural Tasks. In: Proc. Affective Dialogue Systems (ADS), Series: Lecture Notes in Computer Science, Springer, Berlin/Heidelberg, pp. 244–253.

Kopp, Stefan, Gesellensetter, Lars, Krämer, Nicole C., Wachsmuth, Ipke, 2005. A conversational agent as museum guide: design and evaluation of a real-world application. In: Proc. Intelligent Virtual Agents (IVA), Series: Lecture Notes in Artificial Intelligence 3661, Springer, Berlin/Heidelberg, pp. 329-343.

Larsson, Staffan, 2003. Interactive Communication Management in an Issue-based Dialogue System. In: Proc. DiaBruck, Workshop on the Semantics and Pragmatics of Dialogue. Universität des Saarlandes, Saarbrücken, Germany, pp.75–83.

Lohse, Manja, Hanheide, Marc, Pitsch, Karola, Rohlfing, Katharina J., Sagerer, Gerhard, 2009. Improving HRI design applying Systemic Interaction Analysis (SInA). In Proc. Human-Robot Interaction (HRI), 2009.

Lohse, Manja, Hegel, Frank, Wrede, Britta, 2008a. Domestic Applications for social robots - a user study on appearance and function. Journal of Physical Agents, 2 (2), 21-32.

Lohse, Manja, Rohlfing, Katharina, Wrede, Britta, Sagerer, Gerhard, 2008b. Try Something Else! When Users Change Their Discursive Behavior in Human-Robot Interaction. In: Proc. Robotics and Automation (ICRA), pp. 3481-3486.

Muhl, Claudia, Nagai, Yukie, 2007. Does Disturbance Discourage People from Communicating with a Robot? In: Proc. Robot and Human Interactive Communication (RO-MAN), Jeju, Korea, pp. 1137-1142.

Poggi, I., Pelachaud, C., de Rosis, F., Carofiglio, V., De Carolis, B., 2005. GRETA. A Believable Embodied Conversational Agent. In O. Stock, & M. Zancarano

(Eds.), Multimodal Intelligent Information Presentation. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 3-25.

Takeuchi, Masashi, Kitaoka, Norihide, Nakagawa, Seiichi, 2004. Timing detection for realtime dialog systems using prosodic and linguistic information. In: Proc Speech Prosody (SP), pp. 529–532.

Tannen, Deborah, 1984. Conversational style: Analyzing talk among friends. Ablex, Norwood, NJ.

Thórisson, Kristinn Runar, 1996. Communicative Humanoids: A Computational Model of Psychosocial Dialogue Skills. Doctoral thesis, MIT.

Vilhjálmsson, Hannes Högni, Cassell, Justine, 1998. BodyChat: Autonomous Communicative Behaviors in Avatars. In: Proc. Autonomous Agents (Agents), pp. 269–276.

Vollmer, Anna-Lisa, Lohan, Katrin Solveig, Fischer, Kerstin, Nagai, Yukie, Pitsch, Karola, Fritsch, Jannik, Rohlfing, Katharina J. Wrede, Britta, 2009 (submitted). People Modify Their Tutoring Behavior in Robot-Directed Interaction for Action Learning. Submitted to: ICDL 2009.

Ward, Nigel, Tsukahara, Wataru, 2000. Prosodic features which cue back-channel responses in English and Japanese. Journal of Pragmatics, 32, 1177–1207.

Wiener, Norbert, 1948. Cybernetics and Control and Communication in the Animal and the Machine. MIT Press, Boston, MA.

Yngve, Victor H., 1970. On getting a word in edgewise. In Papers from the 6th Regional Meeting of the Chicago Linguistics Society. University of Chicago, Chicago, pp. 567–578.