"She is just stupid"

Analyzing user-agent interactions in emotional game situations

Rehm, Matthias

Published in:
Interacting with Computers

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com
ScienceDirect
Interacting with Computers 20 (2008) 311–325
Interacting with Computers
www.elsevier.com/locate/intcom

"She is just stupid"—Analyzing user–agent interactions in emotional game situations

Matthias Rehm *

Multimedia Concepts and Applications, University of Augsburg, Eichleitnerstr. 30, D-86159 Augsburg, Germany

Available online 13 February 2008

Abstract

A multiplayer dice game was realized which is played by two users and one embodied conversational agent. During the game, the players have to lie to each other to win the game and the longer the game commences the more probable it is that someone is lying, which creates highly emotional situations. We ran a number of evaluation studies with the system. The specific setting allows us to compare user–user interactions directly with user–agent interactions in the same game. So far, the users' gaze behavior and the users' verbal behavior towards one another and towards the agent have been analyzed. Gaze and verbal behavior towards the agent partly resembles patterns found in the literature for human–human interactions, partly the behavior deviates from these observations and could be interpreted as rude or impolite like continuous staring, insulting, or talking about the agent. For most of these seemingly abusive behaviors, a more thorough analysis reveals that they are either acceptable or present some interesting insights for improving the interaction design between users and embodied conversational agents.

Keywords: Embodied conversational agents; Social interaction; Multiuser interaction

1. Introduction

Apart from their conversational skills, the non-verbal behavior and the appearance of embodied conversational agents become more and more realistic. Thus, embodied conversational agents offer great promise to more natural interaction in social settings like tutoring or gaming. Prominent examples of such agent systems include the Steve Agent (Rickel and Johnson, 1999), the real estate agent REA (Cassell et al., 2000), the GRETA Medical Advisor (Pelachaud et al., 2002), the agent MAX (Kopp et al., 2003), the virtual patient (Hubal and Day, 2006) or the tactical language training agent (Core et al., 2006).

Despite these successful examples, virtual agents are still research prototypes and seldom tested with real users in real settings. Thus, not much is known about the acceptance or behavior towards such artifacts that goes beyond small-scale user studies in laboratory settings.

Reeves and Nass (1996) claim that every type of media and especially new media like virtual agents is treated as a social communication partner in one on one interactions. They call this the CASA paradigm—Computers Are Social Actors—and have shown in a number of experiments derived from typical human–human interaction studies that people indeed seem to treat computers and other media like they treat real people. For instance, to attribute technical expertise to a system it matters if a male or female synthesized voice is used. Although the CASA paradigm seems to hold true for certain aspects of interactions with virtual characters, results of actual user studies are not conclusive and sometimes even contradict each other. Krämer (2005) gives an excellent overview of a number of different studies. To give an example, it was shown that the use of an embodied conversational agent can have both positive and negative effects on solving a given task.

All of these studies and all of the above mentioned systems focus on dyadic interactions between one user and one agent. If we turn to communications with more than two interactions partners, we find systems where one user

* Tel.: +49 821 598 2343; fax: +49 821 598 2349.
E-mail address: rehm@informatik.uni-augsburg.de

0953-5438/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.intcom.2008.02.005

engages in the interaction with several agents. A number of approaches to such multiparty conversations have been inspired by research on interactive drama that aims at integrating a user in a scenario—either as an audience member or an active participant. An example includes Avatar Arena (Rist et al., 2003), where the agents perform a presentation for the user interacting among each other. An interesting feature of Avatar Arena is the simulation of listener as well as speaker behaviors based on empirical studies of human–human conversations. Traum and Rickel (2002) have addressed the issue of automatically generated multiparty dialogues in immersive virtual environments. The characters are based on the Steve architecture which has been enhanced by a multi-modal dialogue model to handle turn taking in such a challenging scenario. The VicTec system (e.g., Paiva et al., 2004) realizes a multi-agent learning environment to teach kids strategies against bullying relying on a Forum Theatre metaphor. The user is able to interact with one of the agents and suggest plans of action that will influence the storyline. In the NICE fairy tale game (Gustafson et al., 2004), a user can enter the fairy tale world of H.C. Andersen where she may meet three different types of agents like the helper agent that accompanies her through the world, suggesting lines of action.

Hardly any work so far has been conducted on the realization of scenarios with multiple users and a synthetic agent. An exception includes the work by Isbister et al. (2000) who concentrate on social interactions between several humans in a video chat environment which is supported by a so-called Helper Agent. In contrast to Isbister et al., we focus on a game scenario in which the agent does not appear in the role of a moderator, but takes on a similar role as the human users. The questions arise, if users regard an artificial agent as an equal conversational partner worthy of being attended to even if there are other human interlocutors around and what kind of behavior the users will show towards the agent.

In this article, we are still not moving outside the laboratory setting but our scenario differs in an important aspect from the standard user evaluation. Instead of one user and one agent who engage in a task-oriented dyadic interaction, our game scenario features two users and an agent who play a game of dice together. This offers completely new ways of evaluating the interaction with virtual agents. On the one hand this setting allows us to directly compare the social behavior of the users towards each other with their social behavior towards the virtual characters at the same time in the same situation. On the other hand, users are free to talk with one another which gives us additional insights into what they think about the agent and how they interpret its behavior.

In the next section, the game environment is shortly introduced (Section 2). Section 3 gives an overview of the verbal and nonverbal behavior that is examined in the evaluation studies. Section 4 then deals with the evaluation of gaze behavior which is followed by an evaluation of the users' verbal behavior (Section 5). After a general discussion (Section 6) the article closes with a conclusion (Section 7).

2. Gamble: a multiuser environment for emotional interactions

2.1. The game

Gamble is a game of dice that is played by at least three players. The players have to cast the dice and then they have to announce a result with the constraint that the result must be higher than the one announced by the previous player. The next player can either believe the announcement and cast the dice himself or he does not believe it, the actual result is revealed and depending on the truth of the announcement the winner of the round is declared. Let us look at a short example. Because the interpretation of the cast is a bit complicated, Table 1 gives an overview of possible casts. The cast is interpreted in the following way: the higher digit always represents the first part of the cast. Thus, a 5 and a 2 correspond to a 52. Two equal digits (11,...,66) have a higher value than the other casts, the highest cast is a 21. Now let's assume that player 1 is on turn. He casts the dice and then inspects the dice without permitting the other players to have a look. Player 1 has to announce his cast with the constraint that he has to say a higher number than the previous player. For instance, if the dice show a 52, but the previous player already announced a 61, player 1 has to say at least 62. Now player 2 has to decide whether to believe the other player's claim. In this case, he has to cast next. Otherwise, the dice are shown and if player 1 lied about his result he has lost this round and has to start a new one.

For the Gamble system, one player in this multiplayer game was substituted by a virtual human. The inherent complexity of such multiuser scenarios can be controlled due to the round-based character of the game. At every time of the game, exactly one player has the floor. And he can either announce a result or rate the announcement of another player. Thus, the game can be modeled by a finite state transducer with just three states (see Fig. 1). This transducer completely describes the game flow and is utilized by the central control component of the Gamble system, the game server.

Table 1
Order of casts and probability that the announcement for a given cast is true

Cast	Prob.	Cast	Prob.	Cast	Prob.
31	1	53	0.61	11	0.22
32	0.94	54	0.56	22	0.19
41	0.89	61	0.5	33	0.17
42	0.83	62	0.44	44	0.14
43	0.78	63	0.39	55	0.11
51	0.72	64	0.33	66	0.08
52	0.67	65	0.28	21	0.06

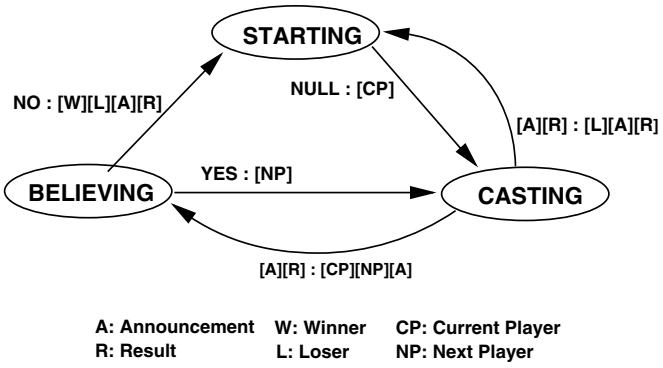

Fig. 1. Overview of game states in Gamble along with in- and output symbols.

2.2. Emotional interactions

Emotional models have increasingly become the center of interest in agent research over the last decade. Partly, this is attributable to the fact that one's own emotions play a crucial role in decision making (see e.g. Damasio, 1995), partly to the fact that emotional information can easily be expressed by facial expressions, and partly to the fact that it becomes feasible to register the user's emotional state and thus integrate it in the agent's decision making process (see Picard, 1997). Here we will focus primarily on the first reason, i.e. emotional models for influencing the behavior of a virtual character.

One approach that has proven successful in this regard and is widely used for implementing an agent's emotional behavior is the OCC-model. This model was developed by Ortony et al. (1988) for describing the cognitive appraisal processes that result in emotional reactions to perceivable stimuli. The model was originally designed to capture the reactions to other people's actions and emotions. To this end, a hierarchical taxonomy of 22 emotions is defined that places the different emotions in 4 non-overlapping groups like *Attraction Reactions* or *Event Reactions*. To give an example of an emotion from the group event reactions let's say you have to catch a train at eight o'clock, you managed to get the bus at half past seven but the bus to the station is stuck in the morning traffic jam. This event is not compatible with your goal of catching the train and consequently results in the emotion of distress. Well-being emotions are triggered by events and are constituted of the two emotions joy (event is compatible with goals) and distress (event is not compatible with goals).

In the large number of implementations that exist, the model is also employed to calculate the agent's emotionally influenced behavior. The wide use of this model is attributable to a great degree to the fact that it is computationally straightforward. Examples of OCC-implementations include the EmotionEngine (Gebhard et al., 2004), Émile (Gratch, 2000), or the Affective Agent Architecture of FearNot! (Aylett et al., 2006). In the FearNot! agent architecture, the OCC-model serves the two functions of appraisal and coping. The appraisal process calculates the emotional state that results from a given stimuli for the agent. The coping process then models the emotional influence on the action selection. It is best described by an iterative execution of the appraisal process, this time not on external stimuli but on internally generated plans.

A different approach that is widely employed makes use of a dimensional model of emotions (Lang, 1995). Such a model is often used for emotion recognition as well as emotion generation (Ball, 2002; Kim et al., 2005; Prendinger et al., 2004). Typically, emotions are defined by the two dimensions of arousal and valence. Emotional states are then represented on two dimensions, one denoting the arousal of the accompanying emotion, the other denoting its valence on a positive/negative axis. Ball (2002) makes use of a Bayesian network to model such a dimensional approach.[1] To this end, arousal and valence are connected to different types of observable behavior like word types, speech features, facial expressions or gestures. Prendinger et al. (2004) describe a similar system that makes use of physiological sensors to assess the user's emotional state in terms of valence and arousal and then uses a Bayesian network to calculate an appropriate emphatic reaction of the agent.

Following these ideas, a two-dimensional model of emotions was integrated in Gamble to render the agent's behavior more interesting and believable on the one hand, and to prepare the integration of emotion recognition techniques on the other hand. The agent's emotional state is influenced by its game success and by its personality traits. Catching another player lying, getting away with a lie or being falsely accused of a lie and thus winning the round constitute a positive emotional influence. On the other hand, falsely accusing another player or being caught lying constitutes a negative emotional influence. The agent's emotional state is also influenced by its so-called personality traits. Instead of using a sophisticated personality model like the five factor inventory (e.g. McCrae and Costa, 1991), we take the dimensional model into account directly. Modulator values for valence and arousal can be specified. These modulators allow the agent different appraisals of a given situation. A high modulator value for valence is interpreted as a highly emotion driven decision process changing fast between positive and negative evaluation of a situation whereas a low modulator value results in a more rational decision. The arousal modulator on the other hand determines how capricious the agent reacts. A high value of the arousal modulator results in a fast increase of the arousal level in a given situation whereas a low value slows the increase down making the agent more phlegmatic. The emotional state of the agent cannot be directly observed but influences

[1] Ball makes use of a three-dimensional model that goes back as far as Wundt (1896) who specified a third dimension that could be called dominance to distinguish e.g. between emotions like fear and anger that have both a negative valence and a high level of arousal.

the behavior selection process, and can thus be perceived through the agent's facial expressions, gesture use, and game play behavior. For instance, if a very angry agent is caught lying, it might very likely start to insult the player who caught the agent. Although such a behavior is not generally acceptable, it is very appropriate in this highly emotional situation and well received by the users who experienced it so far (see Section 5). A happy agent on the other hand, might accept announcements from a player that would have been checked, if a purely rational decision would have been made.

2.3. Agent features

To realize the game, the Greta agent system (Pelachaud et al., 2002) was employed. It features a full body agent which is MPEG-4 compliant and thus allows for realistic facial expressions and gesture animation. The standard speech synthesis component was exchanged against the Loquendo Speech Engine which allows for almost natural speech quality.

2.3.1. Behavior control

The behavior control of the agent takes into account objective information like the probability of an announcement and subjective information from the agent's emotional model. To decide for the agent's context- and situation-specific verbal and non-verbal behaviors, a Bayesian network is deployed. Depending on the evidence available, the network calculates probabilities for possible actions. A turn in the game can roughly be divided into two phases: rating and announcing. First, the announcement of the previous player has to be rated. This decision is based upon (i) the agent's current emotional state (see Section 2.2), (ii) the probability of beating the announcement, and (iii) the number of times that the previous player was caught lying. If the agent is in a positive emotional state it is more likely that it will accept even a high announcement. With each announcement, a certain probability is associated which describes how easily the announcement can be beaten (see Table 1). A 31 for example is the lowest possible announcement and can always be accepted. A 44 on the other hand leaves only three possible results the agent has to achieve next and is thus not easily accepted. Additional information is gained from the agent's memory of the player's previous failures. If a player was caught repeatedly, the agent will treat his announcement with more care.

If the agent believes the previous player or has falsely accused him of lying, it has to cast the dice next and announce a result. The announcement is based upon (i) the agent's cast, (ii) the probability to beat the necessary announcement, (iii) the number of times the agent was already caught lying, and (iv) the agent's emotional state. If the result is higher than the previous player's announcement, there is nothing much to decide. The agent can announce the true result. In case the cast is lower, the agent is forced to lie about the result. In this case the probability which is needed to beat the necessary announcement is taken into account. If the necessary announcement is low, e.g. 43, this probability is high, otherwise it is low, ensuring that the agent is not gambling to high. If the agent was repeatedly caught lying by the next player this will result in a more cautious announcement. A positive emotional state of an agent at last might result in increasing the risk for the announcement.

To sum up, the agent's behavior depends on a mix of objective and subjective features ensuring a consistent but nevertheless not totally predictable behavior.

2.3.2. Gestures

The Greta agent already comes with a standard inventory of gestures that are tied to discourse functions. For the use in Gamble, the gestural repertoire was extended by German emblems.

Following the Berlin Lexicon of German Everyday Gestures (BLAG, 2007), 30 different gestures were specified for the use in Gamble. The rationale for choosing those gestures was threefold. First, they are well documented including the use, as well as the form and the meaning of the gestures. Second, they are clearly identifiable by German native speakers. 90.5% of the generated gestures are correctly classified by subjects. Third, at least half of the documented gestures have a clearly emotion related meaning, e.g., *Waving a hand in front of one's eyes* thereby indicating that something is stupid, *Indicating to one's wrist* which can be interpreted as "hurry up", or *Holding the hand as an extension of the nose*, a clear sign of gloating. This makes them suitable for the use in the game scenario where highly emotional situations arise that ask for the display of appropriate reactions. The agent's gestures thus reveal its emotional appraisal of the situation allowing the other participants to form an impression of the agent's emotional state.

2.3.3. Facial expressions

Although Greta already offers an extensive library of facial expressions, this had to be augmented by expressions accompanying lies to prevent the agent from always showing the perfect poker face. According to Ekman (2001), there are at least four ways in which facial expressions may vary if they accompany lies and deceptions: micro-expressions, masks, timing, and asymmetry.

- Micro-expressions: A false emotion is displayed, but the felt emotion is unconsciously expressed for the fraction of a second. The detection of such micro-expressions is possible only for a trained observer.
- Masks: The felt emotion (e.g., fear) is masked by a different facial expression, in general by a smile. Because we are not able to control all of our facial muscles, such a masking smile is always in some way deficient and reveals at least in part the original emotion.

- Timing: Facial expressions accompanying true emotions do not last for a very long time. Thus, the longer an expression lasts the more likely it is that it is accompanying a lie.
- Asymmetry: Voluntarily created facial expressions as they occur during lying and deceiving tend to be displayed in an asymmetrical way, i.e., there is more activity on one side of the face than on the other.

For our purposes, masks and asymmetry seemed to be the most promising expressions. We concentrated on masking smiles for disgust, sadness, anger, and fear. Different masks are deficient in several aspects. For instance, we considered masks where the eyebrows are still frowning in anger, but the mouth displays a normal smile as well as masks where the frown is not very articulated and there is only a small smile. Different degrees of masking are combined with different degrees of asymmetry of the facial displays resulting in 32 possible facial expressions. A pilot study showed that users were able to notice the differences between genuine and faked smiles (Rehm and André, 2005a).

2.4. Interaction design

Two different interaction devices have been developed for the Gamble system (Figs. 2 and 3). The first version features interaction via a mobile handheld device (PDA) for each player. Casting the dice, announcing the result, and rating the announcement of another player is done via a graphical user interface on the display of the handheld device. The second version features more natural interaction by speech recognition and a tangible interaction device which resembles a cup of dice but is equipped with a webcam to analyze the actual cast and decide on the game progress (for a detailed discussion of the two version see Rehm and André, 2007).

3. Verbal and nonverbal interaction behavior

In contrast to the more traditional dyadic interactions between one user and one agent, our setting allows us to directly compare user–user interactions with user–agent interactions in the same situation at the same time. This gives us some obvious advantages over the standard evaluation, where behavior is compared to the human-only case either with findings from the literature or with previous video recordings of similar interactions between human interlocutors. Because an embodied conversational agent is not another human, interaction can only resemble human interactions but will never be exactly the same.

The general research question we address in our evaluation is "How do people treat the virtual interaction partner, if another real human interaction partner is available?" To play the game, it is not really necessary to socially interact with the virtual character. Our virtual agents are (at least for the time being) always deficient in one or the other way, be it in a limited animation repertoire, be it in limited conversational capacities, or be it in limited perceptual abilities. Thus, in a setting where another human partner is available we can be sure that the behavior towards the agent will deviate from behavior that is "natural" for human–human interactions.

To address this research question, we concentrate on two types of behavior which will be addressed in the next two subsections: (i) gaze behavior and (ii) verbal behavior.

3.1. Gaze behavior

According to Kendon (1967), we can distinguish between at least four functions of seeking or avoiding to look at the partner in dyadic interactions: (i) to provide visual feedback, (ii) to regulate the flow of conversation, (iii) to communicate emotions and relationships, (iv) to improve concentration by restriction of visual input. Kendon showed that speakers tend to look away at the beginning of an utterance and turn their attention towards the conversational partner at the end of an utterance. Goodwin (1981) found the same effect while analyzing turn beginnings and explained that this behavior gives the listener enough time to concentrate on the speaker and direct his gaze towards him. If the listener does not follow this pattern, irritations arise and repair mechanisms like restarts of the turn are initiated by the speaker. Regarding the lis-

Fig. 2. Users interacting with the agent in Gamble v1.0 (left) and v2.0 (right).

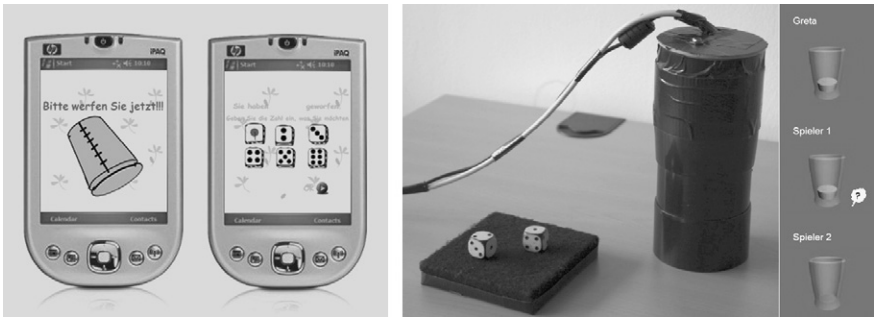

Fig. 3. User interaction via handheld device in Gamble v1.0 (left) and via CamCup, speech recognition and public GUI in Gamble v2.0 (right).

tener, Argyle and Cook (1976) showed that people look nearly twice as much while listening (75%) than while speaking (41%).

Compared to dyadic conversations, we know little about gaze behavior in multiparty interactions. Vertegaal et al. (2001) describe a study of the gaze behavior in a four-party interaction. Subjects looked about 7 times more at the individual they listened to (62%) than at others (9%). They looked about three times more at the individual they spoke to (40%) than at others (12%). In accordance with Sidner et al. (2004) or Nakano et al. (2003), they conclude that gaze is an excellent predictor of conversational attention in multiparty conversations. Among other things, Vertegaal et al. also showed that (i) people look more at the person they speak or listen to than at others, and that (ii) listeners in a group can still see they are being addressed. Each person still receives, 1.7 times more gaze than could be expected had she not been addressed.

The rationale for concentrating on gaze behavior is that we have a setting where a user is interacting both with another human and an embodied conversational agent at the same time. Thus, the user literally has the choice between a natural interaction partner and a moving image on the wall. Our hypothesis is that the user's gaze behavior will roughly resemble the behavior found for human–human interaction (following the CASA paradigm) but will concentrate more on the other human interaction partner, as it is not necessary for the game flow to pay a lot of attention to the agent.

3.2. Verbal behavior

Brown and Levinson (1987) show that politeness is one of the most fundamental verbal behaviors that is found in every language. They develop a system of politeness strategies that is based on Goffman's notion of face. To keep one's own and the interlocutor's face in a communication, it is necessary to mitigate what they call face threats, attacks on face aspects of the interlocutor. Brown and Levinson distinguish between positive and negative face aspects. Positive face denotes the individual's desire to be evaluated positively whereas negative face is concerned with the individual's desire to act free from impositions. According to the two types of face they also introduce positive and negative politeness strategies.

Using appropriate politeness strategies exemplifies that the speaker cares for the addressee's face and thus accepts her as an interaction partner. But is the lack of such strategies identical with the opposite conclusion? To answer this question, we have to look closer into the opposing notion of impoliteness, which is not simply defined by a lack of politeness. Unfortunately, Brown and Levinson only deal with politeness strategies, thus defining what is not impolite behavior. Some have argued that Brown and Levinson's direct bald on record strategy which does not care about mitigating the face threat is a form of impoliteness. But as Bousfield (2003) shows, even a bald on record strategy is concerned with keeping the addressee's face and can only be applied under special circumstances like an immediate danger whereas impoliteness is explicitly concerned with not keeping the face.

To handle this problem Bousfield follows Culpeper's, 1996) definition of positive and negative impoliteness strategies. Culpeper embraces Brown and Levinson's distinction of different face aspects but acknowledges the need for distinguishing different ways of explicitely not respecting the addressee's face needs. The following overview is based on Bousfield (2003):

- Positive impoliteness
  - Ignore or fail to attend to addressee's interests, wants, needs, goods, etc.
  - Exclude the other from activity.
  - Disassociate from the other. Deny common ground, or association.
  - Be disinterested, unconcerned, unsympathetic.
  - Use inappropriate identity markers.
  - Use obscure or secretive language.
  - Seek disagreement—sensitive topics or just disagree outright (act as "Devil's advocate").
  - Avoid agreement—avoid agreeing with addressee's position (whether speaker actually does or not).
  - Make the other feel uncomfortable.
  - Use taboo language—swear, be abusive, express strong views opposed to addressee's.
  - Call addressee names—use derogatory nominations.

- Negative impoliteness
  - Frighten—instil a belief that action detrimental to other will occur.
  - Condescend, scorn or ridicule—emphasise own power, use diminutives to other (or other's position), be contemptuous, belittle, do not take addressee seriously.
  - Invade the other's space—literally (positioning closer than relationship permits) or metaphorically (ask for intimate information given the relationship).
  - Explicitly associate addressee with negative aspect—personalise, use pronouns, "I" and "you".
  - Put addressee's indebtedness on record.
  - Hinder—physically (block passage), conversationally (deny turn, interrupt).

Similar to the politeness strategies, the use of impoliteness strategies emphasizes that the speaker interprets the addressee as someone (something) that has face aspects that can be threatened. Otherwise, it would not be worth the while going to the length of applying such communicative strategies. Thus, the use of such strategies with the virtual game partner in Gamble can be interpreted as accepting this agent as a full-blown communicative partner. Does refraining from the use of such strategies—either polite or impolite—then symbolizes that the agent is not seen as worth while to interact with in a communicative appropriate way? This is not necessarily true.

If the users employ appropriate strategies, this would exemplify Reeves and Nass's (1996) proposed CASA paradigm and strengthen the view that agents are really seen as something similar to a human communication partner. But as we have argued above, our agents are always deficient in some ways but nevertheless they have features which make it easy to anthropomorphize the agent but still keep it in a different category from human interlocutors. Bergmann (1988) analyses interaction with and about pets and gives some interesting findings on how they are similar yet at the same time different from interactions with human interlocutors. His analysis gives us a clue as to how agents might be interpreted by the users. Bergmann shows that utterances directed towards a pet are soon changed into utterances about the pet, which seems to be the dominant form of interaction with/about pets. Bergmann distinguishes four different ways of talking about pets:

(1) Anecdotes narrate real events in which the pet takes a major role.
(2) Interpretation of behavior: The pets behavior is not always directly understandable and has to be interpreted in a "human" way.
(3) Evaluation: The pet or the pet's behavior is evaluated, again in relation to a "human" standard.
(4) Attribution of a motive: The rationale behind the pet's actions has to be explained, thus a motive is assigned to the pet's behavior.

Thus, his analysis might give us a clue on what happens if users start interacting with technical artifacts like virtual interface agents. But Bergmann gives us an additional insight. Communication with pets turns often in communication about pets if another human interaction partner is available. Thus, the Gamble setting is ideal to evaluate if a similar effect can be seen.

Because the user has the choice between a real human interaction partner and the virtual game partner, our hypothesis is again that the user will concentrate on the second human. It is not necessary for the game flow to address the agent directly or start a verbal interaction with the agent.

4. Evaluating gaze behavior

Two evaluation studies with 20 participants were conducted with the two versions of the Gamble system. 10 pairs played at least two games, which each lasted for at least 10 minutes. Positions were changed between games to ensure that each participant played at least once before and after the agent. Interactions were recorded on video. The video recordings were used to analyze gaze behavior as well as verbal interactions. Twelve students from the computer science and from communication science participated in the first evaluation with Gamble v1.0, eight students from computer science in the second evaluation with Gamble v2.0.

As we have seen above (Section 3.1) gaze behaviors differ whether one is in the role of speaker or listener. Thus, two questions guided the evaluation of gaze behavior in the Gamble studies:

- Do people apply different attentive behavior patterns in multi-party scenarios when talking to an agent as opposed to talking to a human?
- Do people apply different attentive behavior patterns in multi-party scenarios when listening to an agent as opposed to listening to a human?

The following hypotheses result from these questions:

(1) Based on Argyle and Cook (1976) and Vertegaal et al. (2001), we assume that humans spend more time on looking at the agent when listening to it than when talking to it.
(2) Following Kendon (1967), we expect similar behaviors at sentence boundaries as in human-human communication.
(3) Since the communicative skills of the agent are strongly limited, we expect that the user will pay more attention to the other human conversational partner. For instance, the user might not establish frequent gaze contact with the agent since he does not expect it to notice it anyway.

4.1. Results

Starting with some basic statistics for Gamble v1.0, Fig. 4 (left) shows the number of gazes towards each of the given directions. The players looked roughly as often towards the virtual agent (27%) as towards the other human player (30%). Just considering the number of gazes, the agent seems to be as attractive as the other player. The fact that people look slightly more often at the PDA (38%) could be attributed to its use as the interface for casting the dice and indicating belief or disbelief. Analyzing the time users spent gazing in the given directions gives a different picture (see Fig. 5 (left)). The PDA is looked at for more than half of the time (55%) and thus draws too much attention away from the interaction partners. This effect resulted in the development of the CamCup and the integration of the speech recognition engine as unobtrusive interaction devices. For the evaluation of Gamble v2.0, the number and length of gazes towards the interaction devices (CamCup and Microphone) and the new GUI (Board) vs. towards the agent show the opposite direction to the results we obtained for Gamble v1.0 (see Fig. 4 (right) and Fig. 5 (right)). The agent now draws most of the human players' attention both regarding the number of gazes (38%) and the length of gazes (52% of the time).

The total number of gazes and the length of gazes during the game provide a rough impression of the users' attention towards human and virtual interlocutors. A more detailed analysis of the users' gaze behavior during announcements and belief-statements of the other players (human as well as agent) gave some additional insights. The results of the first study (Rehm and André, 2005b) confirmed a number of findings about attentive behaviors in human–human conversation. For instance, our subjects spent more time looking at an individual when listening to it than when talking to it—no matter whether the individual was a human or a virtual agent. Furthermore, the type of the addressee (agent or human) did not have any significant impact on the duration of the speaker's gaze behaviors towards the

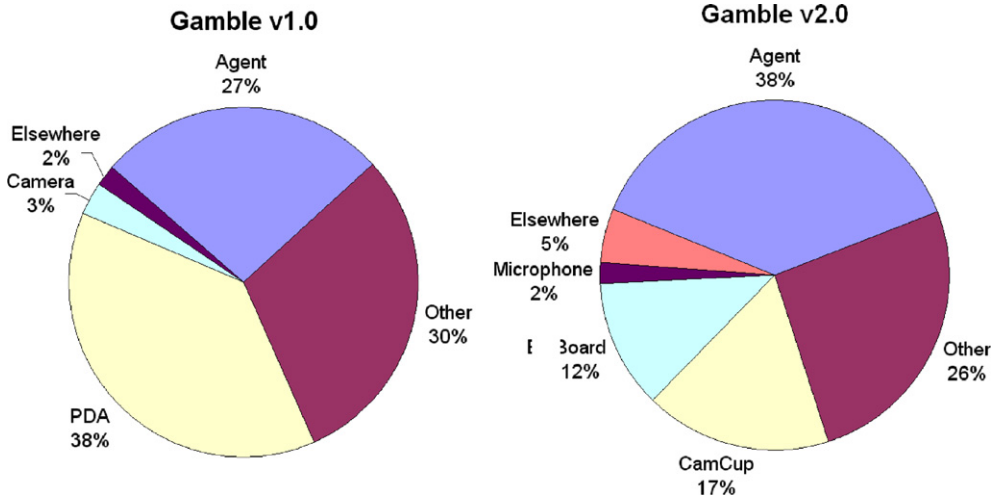

Fig. 4. Number of gazes towards other human and towards agent.

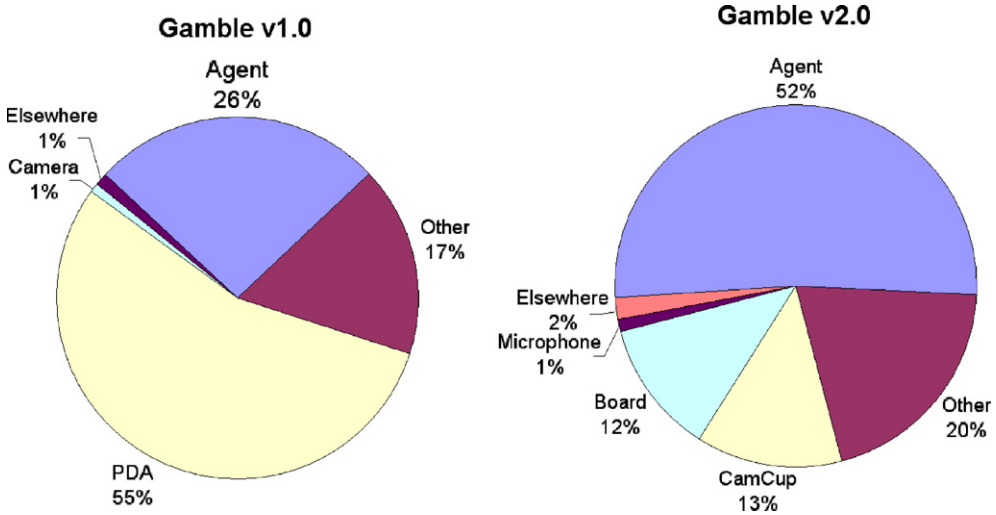

Fig. 5. Length of gazes towards other human and towards agent.

Table 2
Addressee's gaze behavior towards agent or other player

Gaze behavior	Agent	Human	Result
Gamble v1.0	31.75	20.17	$F(1,23) = 23.87^*$
Gamble v2.0	20	14.25	$F(1,7) = 1.92$

* $p < 0.05$.

addressee. Even though the game can be played without paying any attention to the agent's non-verbal communicative behaviors, the users attended to it.

The type of the speaker on the other hand had a significant effect on the addressee at least for the first version of the system (Table 2). This means that the agent is significantly longer looked at when it is speaking compared with the other human player when he is speaking.

As mentioned above, some interesting effects can be found at utterance boundaries. Therefore, we analysed the gaze behavior of speaker and addressee at the beginning and end of utterances, again for the two conditions agent and human. The speaker looks less to the addressee at the beginning and more at the end of an utterances independent on the condition agent or human and thus adheres to the effects found in the literature.

4.2. Discussion

The results show that users adhere in certain aspects to gaze behavior that is described in the literature when they are interacting with the agent. The question arises if the other aspects of users' gaze behavior can be interpreted as rude or impolite. This is not necessarily the case. But we can identify two kinds of gaze behavior that fall into this category.

(1) Ignoring denotes behavior where the interlocutor is not looked at although this would be the acceptable behavior. An example is a speaker that never looks at his addressee during his turn.
(2) Staring denotes behavior where the interlocutor is looked at for a longer time span. It is culturally dependent if this behavior is interpreted as impolite or not (Ting-Toomey, 1999). But at least for the German culture continued staring is interpreted as impolite.

Although our hypothesis was that the users will tend to ignore the agent because they have another human game partner to interact with, the opposite behavior was observed during the evaluation studies. Two different types of staring were visible for the two different versions of the system.

We have seen above, that users adhere to gaze behavior that is described in the literature when they are addressing the agent. This is not always true when they are addressed by the agent (see Table 2). The addressee is following a pattern of gaze behavior that is typical of human-human conversation, in looking towards the speaker at the beginning and away at the end of his utterance. But the agent is attracting more attention than the other human user. For Gamble v1.0 this difference is even significant. This means, the agent is stared at by the user while being addressed by it. During the other time of the game, the PDA draws too much attention to observe the agent continuously. In the evaluation of Gamble v2.0 staring occurs at another time during the game. The attention that was directed towards the PDA in version 1.0 is now directed towards the agent which is observed half of the time of the game (Fig. 5 (right): 52%).

We interpret this effect by the users' need to make sense of the new technical artifact they have encountered. Every participant had already seen different types of agents in lectures they attended. Moreover, they were recruited with a video that showed the Gamble system in action. And they had the possibility to get acquainted with the system before the recording started. Nevertheless, the effect was visible and it did not fade between the different sessions that each pair played. Despite all this, it was the first time that participants actually interacted with an embodied conversational agent. Thus, the close observance of the agent could be attributable to some kind of novelty effect. The two different kinds of staring can be accounted for by the difference in interaction devices. In Gamble v1.0, the PDA drew the attention of the users most of the time. Thus, their only chance to closely observe the agent's behavior was while it was addressing them. With the CamCup and the speech recognition engine in Gamble v2.0, the user's attention was free to closely observe the agent during the game and not only while it was addressing the user.

If this staring behavior is attributable to a novelty effect has to be verified by a long-term study where participants repetitively play the game. If the effect vanishes, there is no need to deal with this type of seemingly rude behavior.

5. Evaluating verbal behavior

The video data collected in the above described evaluation was also used for analyzing the verbal behavior of the users. To this end, the verbal utterances of the users and the agent were transcribed. In contrast to the evaluation of user's gaze behavior, which could be analysed in a quantitative way, the analysis of the transcriptions is mostly done in a qualitative exemplary fashion due to the fact that the users' behavior is not unambiguous. The analysis concentrates on the use of politeness or to be more precise impoliteness strategies on the one hand and on utterances denoting communicative behavior that can be interpreted in line with Bergmann's findings.

5.1. Accepting the agent as a conversational partner

According to Brown and Levinson (1987) the use of the personal pronoun "You" is a positive politeness strategy and signals that the speaker is caring for the face needs of the addressee. Thus, directly addressing the agent in the interaction or directly reacting to the agent's utterances by making use of the second person singular pronoun

"Du" (you) is interpreted as accepting the agent as a conversational partner. The following transcripts give an impression of such interactions.

(1) Addressing the agent
- U: Tja Greta, Fuenferpasch.
  (Well Greta, 55.)
  A: Wird schon stimmen.
  (Will be alright.)
- U: Ich glaub's dir nicht.
  (I don't believe you.)
  A: Du hast verloren.
  (You have lost.)

(2) Reacting to the agent
- A: Na ja, ich glaub' dir mal.
  (Well, I think I believe you.)
  U: Ah so, jetzt glaubst du alles oder?
  (Ah, now you believe everything, don't you?)
- A: Ich habe 61.
  (I have 61.)
  U: Ich glaub's dir.
  (I believe you.)

Thirty-eight percent of utterances that used an address form fall into this category where the agent is directly addressed by the user. This effect was visible in both versions of the system. This means, the agent is sometimes addressed in a very natural way although it is not necessary for the interaction and although in Gamble v1.0 there was not even a speech recognition integrated which could have led to the impression that the agent is able to understand the players' utterances.

The setting of the game makes it difficult to see more politeness strategies at work due to the fact that one has to constantly rate the announcement of others which is either done directly Ich glaub' dir (I believe you) or delivered with mitigating terms like vielleicht (perhaps) or na ja (well) like in Na ja dann glaub ich dir mal (Well then I think I believe you). On the other hand, impoliteness strategies play an important role especially when player's react to having lost a round. In this case, two of the positive impoliteness strategies of Culpeper (1996) can be found frequently in our corpus, namely "using taboo language" and "calling the other names" which can both be classified as insults.

Insults are a special case of verbal interactions in our system. Generally, insults are rarely acceptable in human-human everyday interaction. Such an exception is for instance described by Labov (1972)). He examined "sounding", insult rituals of adolescents in suburban cultures. These rituals have to follow certain rules to be acceptable. In case of Gamble, the game concept is originally a drinking game in Germany, and the fun of the game arises of the interactions between the participants, i.e. how they try to deceive the others, how they react to being caught. When the game gets heated, swearing or insulting are not necessarily seen as an offence but most of the time lead to amused comments of the other players. Thus, although insults can generally be interpreted as impolite behavior, they do not constitute an abuse of the interface agent per se in our scenario. One could even venture out and claim that it is quite contrary a sign of acceptance of the agent, if the user goes to the length of insulting the agent. Obviously, two situations are prototypical for starting to insult the agent. Either the agent caught a player lying about his result or a player did not believe the agent's announcement but was mistaken in this disbelief. In both cases the player loses the round and the agent generally makes a taunting comment on this, often accompanied by an appropriate gesture e.g. a gloating gesture. The following is a collection of insults by the users for the two situations.[2]

(1) Player was caught lying
- U: Bloede Kuh.
  (Stupid cow.)
- U: Verfluchte Scheisse.
  (Damned shit.)
- U: Die kann mich mal.
  (Sod you.)

(2) Player lost because he didn't believe the agent
- U: Ey du pfff bescheisst doch.
  (Ey, you pfff screw me.)
- U: Arschloch.
  (Asshole.)
- U: Oh nein, Greta, Scheisse.
  (Oh no, Greta shit.)

Insulting the other game partners is not a privilege of the human players. The agent's emotional model allows us to simulate an agent that can be really "pissed off" e.g. by having lost continuously. This might then result in insulting the other players. The hardest insult in this case is Du bloedes Arschgesicht (You stupid butthead). This insult is rarely used by the agent. In the two evaluation studies it happened only three times. One of the pairs was quite irritated and decided to ignore this behavior. The other two pairs found this reaction of the agent quite amusing and made comments on the sophisticated interaction capabilities of the agent. Thus, simulating the agent's foul mood in a plausible way, seems to make the agent more believable and acceptable as a game partner.

Generally, insults are rarely acceptable in face to face communication. But in the setting we established with this game of dice, swearing and insulting the others (in a playful fashion) is an acceptable feature of the interaction. Thus,

[2] Translations by author. Comment: a bit difficult to translate the insults into English.

we have an interesting effect here. Although generally we cannot accept that agents insult the user or vice versa, because it is rude, impolite and will surely annoy the user, it is a useful feature here to make the system more believable and acceptable as a game mate. Thus, users that start insulting the agent are really accepting it as a full-fledged game partner that is even able to surprise them (otherwise there would not be a need to get angry about one's own stupidity or the agent's good strategy (and/or luck).

Apart from the above mentioned strategies, the negative impoliteness strategy of "ridiculing the addressee" can be found in the users' interactions with the agent. Most of the agent's phrases were modeled after observations of real players of the game and are very suitable for the game. Now the effect occurs, that some of the more catchy phrases of the agent like `Immer dieses Misstrauen` (Always this mistrust), `Was mach' ich jetzt nur?` (What shall I do now?) or `Da muss ich erstmal ueberlegen` (I have to think about it) were imitated by the human players in corresponding situations. This imitation is not only an adoption of catchy phrases but it is meant as a means of mocking the agent because the users also imitated how the agent spoke the phrases.

The Gamble setting allows us to compare the user's verbal behavior towards the agent with the verbal behavior towards the other player. In case of the address forms, we observed that in roughly half of the cases the second person singular pronoun is used. In the rest of the utterances no address form at all is used like `Glaub' ich` (I believe it) or `Okay` to signal believe of an announcement. The indirect address form which was used with the agent in 62% of the utterances that featured an address form (see next two sections) was never employed when the users talked with each other. The number of insults as defined above by the two positive impoliteness strategies was roughly the same for the human–human and the human–agent case and thus indicates that it is indeed a sign of accepting the agent as a communication partner. The negative politeness strategy of ridiculing the other on the other hand was never used between human players. This indeed seems to be a case of abusive behavior towards the agent that is mocked about due to its limited vocabulary and its slightly odd way of pronunciation.

5.2. Interpreting the agent's behavior

We have seen above that in 38% of the utterances that featured an address form the second person singular pronoun "You" was used. The rest, i.e. 62% of the utterances that used an address form employed the third person singular pronoun "sie" (she) instead, which signifies that the user was talking about the agent and not with the agent. Talking about someone who is actually present in the situation is generally not acceptable. According to Bergmann (1988), even very small children react vehemently if parents violate this "rule" and try to tell a story about the child's behavior to relatives or friends.

This behavior seems only acceptable if the object under discussion is a pet, a baby, or not present. Otherwise it is interpreted as very rude.

Apart from anecdotes, all of the categories described by Bergmann (see Section 3.2) can be found in the Gamble interactions which is exemplified by the following interaction samples.

(1) Interpretation of behavior
- A: So geht das aber nicht. Du hast verloren.
  (What do you think you're doing. You have lost.)
  U: Ich hatte 52. Was will die denn jetzt?
  (I had a 52. What does she want now?)
- A: Ich habe einen Dreier-, aeh Vierer pasch.
  (I have 33, eeh 44.)U: Die luegt nur wenn's sein muss.
  (She is only lying if it is necessary.)
- A: Na, dann werd' ich mal mein Glueck versuchen.
  (Well, then I will try my luck now.)
  U: Jetzt kommt sie wieder mit Dreierpasch oder so.
  (Now she will have a 33 again or something like it.)

(2) Evaluation
- U1: Die ist echt witzig, oder?
  (She is really funny, isn't she?
  U2: Gretchen bekommt die 5 Euro bestimmt.
  (Greta is surely getting the 5 Euro.)
- A: Du hast verloren.
  (You have lost.)U: Ach, die ist doch bloed.
  (Aah, she is just stupid.)

- A: Das nehm' ich dir nicht ab.
  (I don't believe you.)
  U1: Ja ich weiss.
  (Yes I know.)
  U2: Ooh, die zwinkert so schoen.
  (Ooh, she blinks to nicely.)

(3) Attibution of a motive
- A: Aehh, dann sag' ich mal 64.
  (Eeh, I think I say 64.)
  U: Sie will dich verunsichern.
  (She is trying to confuse you.)
- A: 23, ich meine 32.
  (23, I mean 32.)
  U: Du verarscht mich, du machst mich nach.
  (You are fooling me, you are imitating me.)
- A: Jetzt musst du einen Schnaps trinken.
  (Now you have to drink a shot.)
  U: Das Ding, ich wusste nicht, dass das so gemein sein kann.
  (The thing, I didn't know that it can be so mean.)

Table 3 gives an overview on the relative frequencies of the different categories we have identified here. Utterances that are made to evaluate and explain the agent's behavior as defined by Bergmann are the biggest group with 56% (interpretation, evaluation and attribution). This means, users are frequently talking about the agent and trying to make sense of this new technical artifact. This behavior is exclusively seen for the human–agent interactions.

5.3. Ignoring the agent

Additionally to the above categories, we found another recurring form of verbal interaction with the agent, the indirect address. In the sample interactions above, we have already seen the use of the third person singular pronoun "sie" (she). But the utterances were not addressed to the agent. The users tried to interpret the behavior of the agent or evaluated the agent. In the case of what we call an indirect address, the utterance is directly game relevant and directed towards the agent, i.e. either an announcement or a belief statement, but nevertheless the agent is not directly addressed as we have seen in Section 5.1. An example of such a statement is Ich glaub' ihr nicht (I don't believe her). In this example, the user is reacting to the agent's announcement of a result. He could easily have used the direct address form Ich glaub' dir nicht (I don't believe you) or he could have left out the pronoun completely like in Glaub' ich nicht (I do not believe (it)) which would have rendered the utterance totally acceptable. Instead, users frequently opt for the use of this third person singular pronoun and thus signify that the agent is somehow different from the other human player. Of course this kind of address is never used with the other human player.

Table 3
Relative frequency of occurence of verbal categories

Interpretation	Evaluation	Attribution
28%	17%	11%
Indirect Address	Imitation	Insults
22%	5%	17%

6. Discussion

The evaluation studies showed that the agent is both accepted as a game (and conversational) partner by the human players and that the agent is nevertheless treated in a special way and sometimes in a way that would be interpreted as rude if it was directed towards a human interlocutor.

Analyzing the users' gaze and verbal behavior, we can distinguish three different ways of interacting with the agent:

(1) Acceptance represents that the users treat the agent as a game mate similar to the other human interlocutor, adhering to gaze behavior found in human–human interactions, directly addressing it, or reacting to its verbal and non-verbal behavior either in a polite or impolite way.
(2) Interpretations describe situations where the users try to make sense of the agent's behavior either by closely observing the agent or by discussing the behavior with the other player.
(3) Ignoring denotes behavior where the users treat the agent as if it is not part of the setting, addressing it in the third person.

All of these behaviors are always present in the interactions with the agent, i.e. there is no complete acceptance of the agent neither a complete disregard of the agent. The question now arises, how we should deal with this kind of behavior mix from the user. Should we react to his impolite or rude behavior regarding the agent? Or should we accept it as the "natural" way of interacting with the agent? Our answer to this is: both. The behavior we encountered in the evaluation studies has different reasons and serves different purposes. First of all, we have to accept that embodied conversational agents are new technical artifacts that have not been widely introduced to the public yet. Thus, users that are confronted with such agents have first to develop an idea on what the agent is capable of and how it reacts in certain situations. And although embodied agents are designed to simulate natural communication behaviors, agents will never be treated in exactly the same way as a human interlocutor due to the simple fact that they are no humans.

Having said that, the behaviors encountered in the evaluation studies should each be handled in a different way. The impolite behavior like insulting and imitating the agent does not only emphasize that the agent is regarded as a conversational partner whom's face needs can be violated but is in case of the Gamble system even acceptable behavior that is a natural part of the gaming interaction. To make the agent even more believable it should react to an insult with a suitable answer thus acknowledging the playful character of this interaction. If users discuss the agent's features by talking about the agent, this symbolizes the users' need to rationalize about the artifact they are just encountering. In this case, the agent could offer some help in interpreting its behavior instead of just being offended. If on the other hand, utterances that should be addressed to the agent are in fact phrased in a way that implies talking about the agent it might be a reasonable reaction of the agent to make an appropriate comment to this rude behavior because it interrupts the flow of the interaction. Thus, to make the agent even more acceptable as some kind of interaction partner, it would be good to react to such clues.

The same is true for continuous staring of the user. Getting some attention of the user might be acceptable due to the fact that the user is trying to get acquainted with the agent. But staring at the agent for half of the whole interaction time should not be accepted. For the interaction in Gamble it is not so crucial that the other human player does not get as much attention as the agent. But by staring at the agent, the user avoids eye contact with the other human interlocutor. This could get critical in applications that rely on the collaboration of the users.

7. Conclusion

In creating embodied conversational agents as interfaces we allow users to interact with complex systems relying on their everyday communicative abilities. To further this interaction, we try to make our agents ever more lifelike and their verbal and non-verbal behavior as "natural" or to be more precise as human-like as possible. Nevertheless, the above presented results show that the users are of course well aware of the "otherness" of our artifacts and thus come up with new behavior routines for such artifacts. This is not a bad thing but quite to the contrary a necessary step in accepting embodied conversational agents as kind of interaction partners that offer an additional value for the interaction. But as developers of agent systems, we have to take this into account for our system models. This means that we should not only opt for the best comparison with human–human interactions but that we take the specifics of human–agent interactions and the differences to human–human interactions into account to make agents more acceptable, e.g. by letting the agent explain its own behavior if the user starts to rationalize about the agent.

Users anthropomorphize agents and treat them as communication partners but only to a certain degree. This is a weak interpretation of Reeves and Nass's (1996) CASA paradigm (Computers Are Social Actors). Bergmann (1988) makes a similar comment in his work on pets. In his opinion, it is unavoidable to anthropomorphize pets. But he also states that humans often avoid treating them in the same way that they treat other humans. Our argument follows these lines. Agents are of course perceived as anthropomorphic entities, they are designed that way. But they are always in some way deficient be it in not being aware of the context, be it in a restricted vocabulary or in a restricted set of animations, be it in the quality of the speech synthesis. Thus, users will also always know that there is a difference between agents and human interlocutors and they will thus also always treat agents differently to a certain degree.

We have seen in our evaluation of the users' gaze behavior as well as the users' verbal behavior that we can distinguish between three ways of interacting with the agent which are not clearly distinguishable between different users but are mingled together in every interaction. Acceptance denotes that users sometimes accept the agent as a communication partner that is worth talking to (although the agent can only understand some digits and yes/no-variants). Insults, although rude and impolite in general, are another sign of acceptance and attribution of competence towards the agent in this scenario. But mingled with these forms of interaction where the agent is treated as a full-fledged interlocutor, there are also a lot of situations where the agent is treated as an artifact. Is this a problem? We think not, because the users' utterances show that most of the time they talk about the agent, they are trying to figure out what the agent is capable of, to make sense of this somewhat astonishing technical artifact. This gives us information on what kind of features the users attribute to the agent.

Future work on the Gamble system will need to augment the agent's abilities to perceive situational factors like they were apparent in the evaluation studies. A first option is extending the agent's emotional model by a real-time emotion recognition component that analyses the speech

signal and has already successfully been tested in a single user setting (Hegel et al., 2006). An eye tracker could be integrated as another input device which would allow the agent to register the user's gaze behavior and react to continuous staring. Such an integration would mean to move away from a multiple user scenario. But a multiplayer setting is imaginable where one user plays with two agents. Last but not least, a long-term study is planned that will shed light on two questions: (i) To what degree is the users' behavior attributable to some kind of novelty effect? (ii) What kind of stable user–agent interaction patterns emerge after longer interaction periods and how do they differ from human-human patterns of behavior?

Acknowledgements

I am grateful to the editors and three anonymous reviewers for their helpful comments to improve the article. The work described in this article was partly supported by the EU Network of Excellence HUMAINE (http://emotion-research.net).

References

Argyle, M., Cook, M., 1976. Gaze and Mutual Gaze. Cambridge University Press, Cambridge.

Aylett, R., Dias, J., Paiva, A., 2006. An affectively driven planner for synthetic characters. In: Proceedings of ICAPS. AAAI Press, pp. 2–10.

Ball, E., 2002. A Bayesian heart: computer recognition and simulation of emotion. In: Trappl, R., Petta, P., Payr, S. (Eds.), Emotions in Humans and Artifacts. MIT Press, Cambridge, MA, pp. 303–332.

Bergmann, J., 1988. Haustiere als kommunikative Ressourcen, Soziale Welt: Zeitschrift für sozialwissenschaftliche Forschung und Praxis, Sonderband: Kultur und Alltag 8, 299–312.

BLAG, Berliner Lexikon der Alltagsgesten, http://www.ims.uni-stuttgart.de/projekte/nite/BLAG last visited: 31st August 2007.

Bousfield, D., 2003. Impoliteness in Interaction, Ph.D. thesis, Lancaster University.

Brown, P., Levinson, S.C., 1987. Politeness—Some Universals in Language Usage. Cambridge University Press, Cambridge.

Cassell, J., Bickmore, T., Campbell, L., Vilhjalmsson, H., Yan, H., 2000. Designing embodied conversational agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), Embodied Conversational Agents. MIT Press, Cambridge, MA, pp. 29–63.

Core, M., Traum, D., Lane, H.C., Swartout, W., Gratch, J., Lent, M.V., Marsella, S., 2006. Teaching negotiation skills through practice and reflection with virtual humans. Simulation 82 (11), 685–701.

Culpeper, J., 1996. Towards an anatomy of impoliteness. Journal of Pragmatics 25, 349–367.

Damasio, A.R., 1995. Descartes Irrtum, dtv.

Ekman, P., 2001. Telling Lies—Clues to Deceit in the Marketplace, Politics, and Marriage, 3rd ed. Norton and Co. Ltd., New York.

Gebhard, P., Klesen, M., Rist, T., 2004. Coloring multi-character conversations through the expression of emotions. In: André, E. et al. (Eds.), Affective Dialogue Systems (ADS04). Springer, Berlin, Heidelberg, pp. 128–141.

Goodwin, C., 1981. Conversational Organization—Interaction between Speakers and Hearers. Academic Press, New York.

Gratch, J., 2000. Émile: Marshalling passions in training and education. In: Proceedings of the 4th Conference on Autonomous Agents. ACM Press, pp. 325–332.

Gustafson, J., Bell, L., Boye, J., Lindstrm, A., Wiren, M., 2004. The NICE Fairy-tale Game System. In: Proceedings of SIGdial 04.

Hegel, F., Spexard, T., Vogt, T., Horstmann, G., Wrede, B., 2006. Playing a different imitation game: interaction with an empathic android robot. In: Proceedings of the IEEE-RAS International Conference on Humanoid Robots (Humanoids'06), pp. 56–61.

Hubal, R.C., Day, R.S., 2006. Informed consent procedures: an experimental test using a virtual character in a dialog systems training application. Journal of Biomedical Informatics, 532–540.

Isbister, K., Nakanishi, H., Ishida, T., Nass, C., 2000. Helper agent: designing an assistant for human–human interaction in a virtual meeting space. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 57–64.

Kendon, A., 1967. Some functions of gaze direction in social interaction. Acta Psychologica 32, 1–25.

Kim, J., André, E., Rehm, M., Vogt, T., Wagner, J., 2005. Integrating information from speech and physiological signals to achieve emotional sensitivity. In: Proceedings of Interspeech/Eurospeech, pp. 809–812.

Kopp, S., Jung, B., Lessmann, N., Wachsmuth, I., 2003. Max—a multimodal assistant in virtual reality construction. KI—Künstliche Intelligenz 4, 11–17.

Krämer, N.C., 2005. Social communicative effects of a virtual program guide. In: Intelligent Virtual Agents 2005. Springer, Berlin, Heidelberg, pp. 442–453.

Labov, W., 1972. Rules for ritual insults. In: Sudnow, D. (Ed.), Studies in Social Interaction. Free Press, pp. 120–169.

Lang, P.J., 1995. The emotion probe: studies of motivation and attention. American Psychologist 50 (5), 372–385.

McCrae, R.R., Costa, P.T., 1991. The neo personality inventory: using the five-factor model in counseling. Journal of Counseling and Development 69, 367–372.

Nakano, Y.I., Reinstein, G., Stocky, T., Cassell, J., 2003. Towards a model of face-to-face grounding. In: Proceedings of the Association for Computational Linguistics. Sapporo, Japan, pp. 553–561.

Ortony, A., Clore, G.L., Collins, A., 1988. The Cognitive Structure of Emotions. Cambridge University Press, Cambridge.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Woods, S., Zoll, C., 2004. Caring for agents that care: building empathic relations with synthetic agents. In: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems. ACM Press, pp. 194–201.

Pelachaud, C., De Carolis, B., de Rosis, F., Poggi, I., 2002. Embodied contextual agent in information delivering application. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems. ACM Press, pp. 758–765.

Picard, R., 1997. Affective Computing. MIT Press, Cambridge, MA.

Prendinger, H., Dohi, H., Wang, H., Mayer, S., Ishizuka, M., 2004. Empathic embodied interfaces: addressing users' affective state. In: André, E. et al. (Eds.), Affective Dialogue Systems (ADS-04). Springer, Berlin, Heidelberg, pp. 53–64.

Reeves, B., Nass, C., 1996. The Media Equation—How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, Cambridge.

Rehm, M., André, E., 2005a. Catch me if you can—exploring lying agents in social settings. In: Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems. ACM Press, pp. 937–944.

Rehm, M., André, E., 2005b. Where do they look? Gaze behaviors of multiple users interacting with an embodied conversational agent. In: Panayiotopoulos, T., Gratch, J., Aylett, R., Ballin, D., Olivier, P., Rist, T. (Eds.), Intelligent Virtual Agents (IVA'05). Springer, Berlin, Heidelberg, pp. 241–252.

Rehm, M., André, E., 2007. The virtual human in the loop—engaging in social games using tangible interface devices. Journal of Game Development 2 (3), 5–23.

Rickel, J., Johnson, W., 1999. Animated agents for procedural training in virtual reality: perception, cognition and motor control. Applied Artificial Intelligence, 415–448.

Rist, T., Schmitt, M., Pelachaud, C., Bilvi, M., 2003. Towards a simulation of conversations with expressive embodied speakers and

listeners. In: Proceedings of Computer Animation and Social Agents (CASA), pp. 5–10.

Sidner, C.L., Kidd, C.D., Lee, C., Lesh, N., 2004. Where to look: a study of human–robot engagement. In: Proceedings of the Ninth International Conference on Intelligent User Interface, pp. 78–84.

Ting-Toomey, S., 1999. Communicating Across Cultures. The Guilford Press, New York.

Traum, D., Rickel, J., 2002. Embodied agents for multi-party dialogue in immersive virtual worlds. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems. ACM Press, New York, pp. 766–773.

Vertegaal, R., Slagter, R., van der Veer, G., Nijholt, A., 2001. Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In: Proceedings of Eye Gaze Patterns in Conversations: There is More to Conversational Agents Than Meets the Eyes, Proceedings of SIGCHI 2001. Seattle, WA, pp. 301–308.

Wundt, W., 1896. Grundriss der Psychologie. Engelmann, Leipzig.