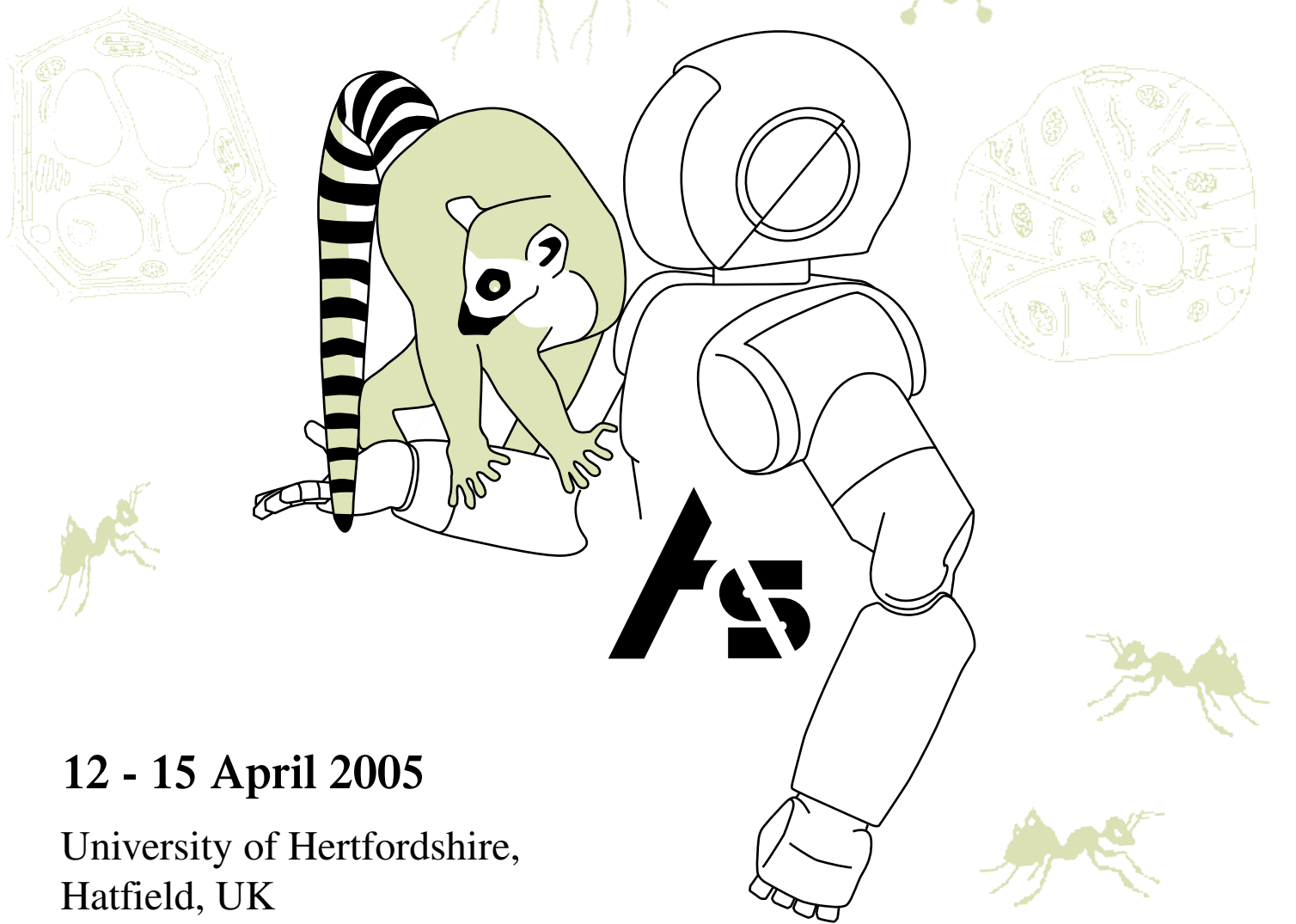AISB'05: Social Intelligence and Interaction
in Animals, Robots and Agents

# Proceedings of the Symposium on Conversational Informatics for Supporting Social Intelligence and Interaction - Situational and Environmental Information Enforcing Involvement in Conversation

**12 - 15 April 2005**

University of Hertfordshire,
Hatfield, UK

**SSAISB 2005 Convention**

AISB    UH    EPSRC
Engineering and Physical Sciences
Research Council

# AISB'05 Convention

*Social Intelligence and Interaction in Animals, Robots and Agents*

12-15 April 2005
University of Hertfordshire, Hatfield, UK

Proceedings of the Symposium on

## Conversational Informatics for Supporting Social Intelligence and Interaction:

Situational and Environmental Information Enforcing
Involvement in Conversation

# Table of Contents

# Informing the Design of Embodied Conversational Agents by Analyzing Multimodal Politeness Behaviors in Human-Human Communication

Matthias Rehm and Elisabeth André

*Multimedia Concepts and Applications
Faculty of Applied Informatics
University of Augsburg
{rehm|andre}@multimedia-werkstatt.org

**Abstract**

In order to build embodied conversational agents that are able to communicate with the user in a more natural manner, the consideration of social aspects seems inevitable. One aspect of social interaction is the use of politeness strategies. In this paper, we report on a corpus study we conducted in order to shed light on the co-occurrence of gestures and verbal politeness strategies in face threatening situations. The results of the study will be used to inform rules for the selection of gestures in an ECA.

## 1 Introduction

Embodied conversational agents (ECAs) are becoming more and more realistic in their appearance and their animations. But supplying an interface agent with a body also poses great challenges to the design of appropriate interactions because the user will expect - at least in part - humanlike verbal and nonverbal conversational behaviors of such an agent. In the long run, it is therefore inevitable to enrich ECAs with social competencies to render their interactions with the user more natural and entertaining. One aspect of social interaction is the use of politeness strategies as they are described in detail in Brown and Levinson's (1987) seminal work. People maintain positive (self image) and negative face (wants and desires), which are continuously threatened during interactions, e.g., by commands or criticism on one's behavior. Such speech acts are called face threatening acts (FTAs). People try to redress or mitigate such undesirable acts, e.g., by referring to the good looks of the addressee before asking her for a favor.

Previous work has concentrated for the most part on the linguistic aspects of FTAs, i.e., on verbal means to deliver and redress FTAs. But FTAs are often multi-modal. Dressing up a threat in a joke usually only works if the speaker shows in his whole appearance (facial expression, body posture) that he is telling a joke. Otherwise the threat might be even more severe than it is (see Fig. 1 for an example).

Due to the sparse literature on the use of non-verbal communicative behaviors of politeness, we collected our own corpus based on staged conversations between humans. To trigger the use of politeness strategies, we had to make sure that the communicative situation was inherently face-threatening for the participants. We therefore decided to record scenarios where an audience had to provide criticism to the speaker. The recorded video material was annotated and analyzed in order to identify frequently occurring combinations of gestures and verbal politeness strategies.

## 2 Related Work

Research on non-verbal communicative behaviors, such as gestures or facial expressions, provides a good impression of the relevance of multi-modal aspects of communication, e.g. (Allwood, 2002), (Kendon, 1986), (Knapp and Hall, 1997), (Pease, 1993), and reveals a bunch of implicit information about the role of gestures and facial expressions in delivering and redressing face threats. However, there is hardly any work that explores the relationship between multi-modal means of communication and face threats. An exception is an empirical study by Trees and Manusov (1998) who found that non-verbal behaviors, such as pleasant facial expressions and more direct body orientation may help to mitigate face threats evoked by criticism. Bavelas et al. (1995) provide a classification of gestures some of which

can be directly mapped onto Brown and Levinson's strategies of politeness. Shared information gestures mark material that is part of the interlocutors common ground. Citing gestures refer to previous contributions of the addressee and aim at conveying the impression that the interlocutors share a common opinion. Elliptical gestures mark incomplete information that the addressee should augment for him- or herself and may take on a similar function as off-record strategies. Seeking agreement gestures directly correspond to Brown and Levinson's approval oriented strategies. Turn open gestures can be regarded as attempts to satisfy the addressee's desire for autonomy. Linguistic means to deliver FTAs have partly become part of the grammar and Bavelas classification of gestures suggests that there might be similar principled and standardized connection between nonverbal means of communication and politeness strategies.

Walker et al. (1997) have presented one of the first approaches to implement politeness strategies as a means to more flexible dialogue control. They summarize the available strategies into four main categories: (1) direct, (2) approval oriented, (3) autonomy oriented, (4) off record. In direct strategies, no redress is used, the speaker just expresses his wishes. Approval oriented strategies are related to the positive face needs of the addressee, using means to approve of her self-image. Autonomy oriented strategies on the other hand are related to the negative face wants of the addressee, trying to take care of her want to act autonomously. Off record strategies at last are the most vague and indirect form to address someone, demanding an active inference on the side of the addressee to understand the speaker. Depending on variables such as social distance and power, and a culture-specific rating of the speech act, a speaker chooses an appropriate strategy to deliver a face threatening act (FTA), e.g. (i) I really enjoyed your talk but you should be more coherent vs. (ii) The talk should be more coherent. In (i) the speaker compliments the addressee on her talk before delivering his critic, thus employing an approval oriented strategy. In (ii), an autonomy oriented strategy is used in impersonalizing the criticism. The speaker neither refers to the addressee nor to himself. Johnson et al. (2004) describe the value of politeness in a tutoring system. Examining the interactions between a real tutor and his students, they came up with a set of templates to generate appropriate utterances depending on the current situation. One interesting modification of the original theory by Brown and Levinson (1987) was to select approval and autonomy oriented strategies based on the type of the expected face threat (and



Figure 1: Critic of the Popidol show wrapping his criticism in a joke.

not just on its weight). André et. al (2004) augmented the model of Brown and Levinson with an emotional layer. The emotion of the addressee as it is observed by the speaker plays a crucial role in determining an appropriate strategy. Bickmore and Cassell (2000) describe how smalltalk is utilized to build up common ground between an embodied conversational agent and the user based on an extension of Brown and Levinson's theory of politeness. Nakano et al. (2003) study how people use non-verbal signals, such as eye gaze and head nods, to provide common ground in the context of direction-giving tasks. Even though their work relies on a sophisticated model of gestural communication, they did not investigate how the use of gestures may help to mitigate the face threat for the user. Porayska-Pomsta and Mellish (2004) make use of Brown and Levinson's model in order to motivate linguistic variations of a natural language generator. Prendinger and Ishizuka (2001) consider Brown and Levinson's social variables distance and power in order to control emotional displays of agents. For instance, if the social distance between an agent and its conversational partner is high, the agent would not show anger to the full extent. This behavior can be interpreted as an attempt to reduce the face threat for the conversational partner.

Summing up it may be said that the implementation of politeness behaviors in an ECA mainly focused on verbal aspects so far.

## 3 The Augsburg SEMMEL corpus

Since there is hardly any research into the multimodal aspects of human politeness strategies, we decided to acquire our own multi-modal corpus for an empirical grounding of the intended system. We explored two alternatives. Our first approach was to rely

on video recordings from the German version of the TV show Popidol (see Fig. 1). In this show, a number of candidates present a song. A jury comments on the performances and the viewers vote for the candidates. After some weeks, the popidol for the season emerges. The advantage of this corpus lies in the fact that the phenomena we are interested in are a major ingredient of the show. Furthermore, the TV personalities were experienced speakers that make use of expressive gestures and facial expressions. On the other hand, their behavior is certainly not representative of ordinary people. Furthermore, the corpus did not provide enough examples of multi-modal politeness behaviors since there was little criticism towards the end of the show and the gestures and facial expressions of the jury were not always visible. Although this corpus gave us interesting insights in the combined use of verbal and non-verbal politeness behavior, the limitations of the corpus only allowed for anecdotal evidence. Thus, we decided to collect a new corpus based on staged scenarios with a group of students.

## 3.1 Collecting the SEMMEL-Corpus

We devised a scenario that forced the participants to use their (unconscious) knowledge of politeness strategies by confronting them with an inherently face threatening situation. Criticizing someone is a prototypical example of such a situation. Therefore, we chose seminar talks with subsequent discussion to provide for a more or less "natural" situation for the participants. The focus was on the criticism given by the audience to the speakers on their performance. Students were divided into two groups: audience and speakers. The speakers were asked to give a five minute talk about one of their hobbies. This topic was chosen to keep the necessary preparatory work for the talk at a minimum and to ensure that the audience had enough knowledge on the topic to easily criticize the speaker.

The initial explanation for this setup that was given to the participants one week before the experiment was our need to collect a corpus of non-verbal communicative behavior. This explanation also accounted for the two cameras we were using, one videotaping the speaker, the other one the audience. The initial explanation was detailed on the day of the experiment. The speakers were informed about the real setup to prevent them from reacting in an unwanted way to the critic or the criticism. The audience was told that we were interested in the reaction of the speaker to (potentially unjustified) criticism. In order to ensure that we would collect enough examples of relevant communicative acts, each member of the audience was instructed to criticize the speaker on three different dimensions and received a list of issues that had to be brought up during the discussion: (i) formal aspects, e.g. too many/too few slides, (ii) content, e.g. snowboarding is far too dangerous, and (iii) personal, e.g. the speaker was to nervous. After the experiment, the participants were informed about the actual objective of the data collection.

12 students in their first and second year participated in this data collection, three male, nine female. Four of them (two male, two female) prepared a talk on their hobby and were criticized by four audience members immediately after their presentation (see Fig. 2). The audience for each talk was constituted randomly from the remaining eight students ensuring that each of them participated two times as an audience member and met one of the other audience members only twice. We tried to hold the social variables distance and power constant and made sure that the speakers and the audience were not from the same year. The resulting SEMMEL-corpus (Strategy Extraction for MultiModal Eca controL) contains 66 different acts of criticism, i.e., 16.5 on average per talk. An act of criticism covers one of the aspects mentioned above and is always delivered with a mix of strategies and co-occuring gestures. Up to now, roughly half of our material has been annotated containing 125 combinations of strategies and gestures.

## 3.2 Annotating the SEMMEL-Corpus

The collected material was annotated using ANVIL (Kipp, 2003). Fig. 2 shows a screenshot of the ANVIL system along with annotations of our corpus. Focusing on the interaction of verbal and non-verbal behavior in the use of politeness strategies, the SEMMEL coding scheme features four main layers:

1. trl: The transliteration, i.e., the words spoken.

2. affective facial expression: Facial expressions that can be labeled with an emotion.

3. gesture: The hand gestures of the speaker visible in the video.

4. strategy: The politeness strategies employed by the speaker.

Focusing on the use of gestures as a non-verbal means to redress face threats, facial expressions are not annotated at the moment. In the coding scheme, facial expressions may be annotated using the affective tags available in APML (Carolis et al., 2002).
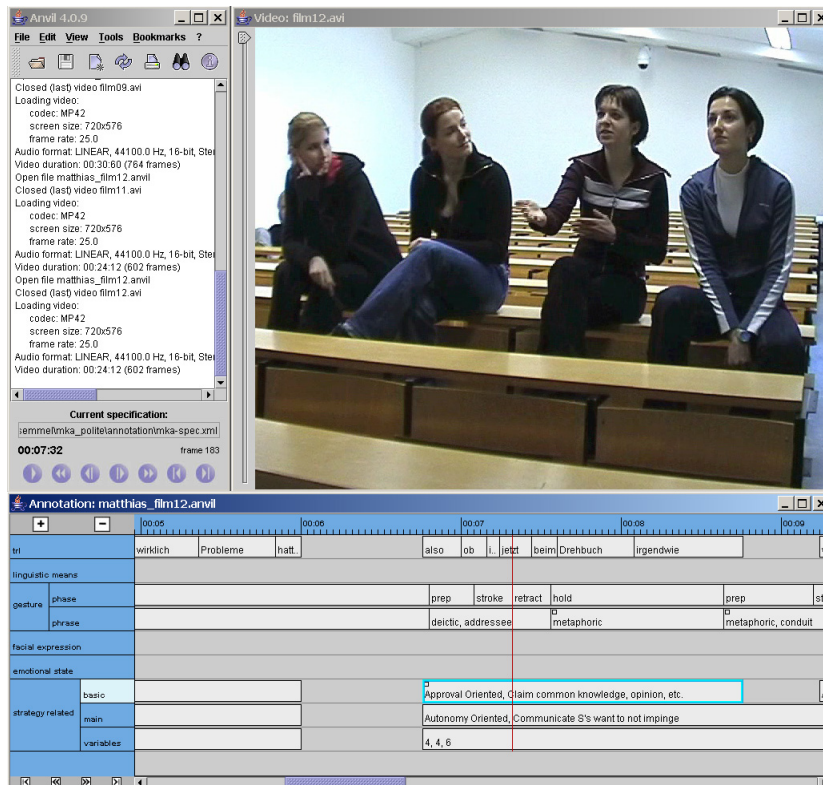
Figure 2: Snapshot from the ANVIL annotation system. Above the video is displayed, below the annotation board.

| Category | Strategy | Verbal means |
|---|---|---|
| Direct | Direct | State the threat directly |
| Approval Oriented | Convey interest | Compliments, intensifying adjectives |
| | Claim in-group membership | Address forms, slang, elliptical utterances |
| | Claim common knowledge | White lies, use of "sort of", "in a way", jokes |
| | Indicate knowledge about wants | State to regard addressee's wants |
| | Claim reflexivity | Inclusive "we", give/ask for reasons |
| | Claim reciprocity | State that addressee's owns speaker a favor |
| | Fulfil wants | State sympathy |
| Autonomy Oriented | Make minimal assumptions | Hedges "I think", "kind of" |
| | Give option not to act | Subjunctive, use of "perhaps" |
| | Minimize threat | Euphemisms, use of "a little", "just" |
| | Communicate want not to impinge | Avoidance of "you" and "I", state threat as general rule |
| | Indebting | |
| Off Record | Violate relevance maxim | Associations, hints |
| | Violate quantity maxim | Exaggerations like "always" |
| | Violate quality maxim | Irony, rhetorical questions |
| | Violate manner maxim | Ambiguity, elliptical utterances |

Table 1: Types of strategies used for coding the SEMMEL corpus and examples of verbal means to realize these strategies.

In this vein, we will get well defined categories of facial expressions that can be used later for generation purposes in a straight forward way.

The coding of gestures follows Kipp's approach (Kipp, 2003) which is based on McNeill's guidelines (McNeill, 1992). Accordingly, two different parts of a gesture are distinguished: the gesture phase and the gesture phrase.

- Track gesture.phase: This is a primary track, which means that it is directly related to the video. Although gestures are mostly co-verbal, i.e., they accompany speech and add additional meaning to it by visualizing aspects of the mentioned referents, only the stroke of the gesture has verbal-nonverbal synchronization constraints. Thus it does not suffice to bind the gesture only to the transliteration layer but to the video itself. The most prominent phases of a gesture are preparation, stroke, and retraction. Generally, the hands are brought from a resting position into the gesture space during preparation. The stroke is the phase of the gesture, that carries/visualizes its meaning. Afterwards, the hands are brought back to a resting position during the retraction phase.

- Track gesture.phrase: The gesture phrase denotes the type of the gesture. It is realized as a secondary track which means it is related to another track of the coding scheme, in this case to gesture.phase. Thus, the gesture phases specify the time dimension of the gesture in regard to the video whereas the gesture phrase gives the interpretation of this specific gesture. McNeill (1992) distinguishes roughly between adaptor, beat, emblem, deictic, iconic, and metaphoric gestures. Adaptors comprise every hand movement to other parts of the body like scratching one's nose. Beats are rhythmic gestures that may emphasize certain propositions made verbally or that link different parts of an utterance. Emblems are gestures that are meaningful in themselves, i.e., without any utterance. An example is the American "OK"-emblem, where the thumb and first finger are in contact at the tips while the other fingers are extended. Deictic gestures identify referents in the gesture space. The referents can be concrete like the addressee or they can be abstract like pointing to the left and the right while uttering the words "the good and the bad". In this case the good and the bad are identified in the gesture space and it becomes possible to refer back to them later on by point-

ing to the corresponding position. Iconic gestures depict spatial or shape-oriented aspects of a referent, e.g., by using two fingers to indicate someone walking while uttering "he went down the street". Metaphoric gestures at last are more difficult in that they visualize abstract concepts by the use of metaphors, e.g. using a box gesture to visualize "a story". This is the conduit metaphor that makes use of the idea of a container in this case a container holding information.

The coding of strategies uses a simplified version of Brown and Levinson's hierarchy distinguishing between seven different approval oriented, five different autonomy oriented, and four different off record strategies (see Table 1).

- Track strategy.basic: Every strategy that is employed by the speaker is coded and bound to the words in the transliteration track that give rise for this interpretation. For each category of strategies (direct, approval oriented, autonomy oriented, off record), the coder has to decide for a specific type (see Table 1).

- Track strategy.main: Because a single utterance contains nearly always a mix of strategies, a track is added for coding the main strategy used in a specific utterance. The same elements as in the basic track are used (see Table 1), but the elements in this track are not bound to the transliteration but to the basic track.

- Track.variables: Brown and Levinson introduce the contextual variables social distance, power relation, and ranking of the imposition to calculate the weight of the face threat that is redressed by the strategy. This track is bound to strategy.main assuming that neither of the variables changes during a single utterance.

### 3.3 Analyzing the SEMMEL corpus

The first part of our analysis concentrated on the distribution of the four basic categories of politeness strategies. Remarkably is the high number of autonomy oriented strategies. From the 125 strategy/gesture combinations, 61% include autonomy oriented strategies, 18% Off record, and 15% Approval oriented strategies. By opting for autonomy oriented strategies, the critics try to leave the choice of action on the side of the addressee. Thus, the criticism is wrapped into some kind of suggestion for the addressee on how to improve the talk. We put this

| Strategy | Freq. |
|---|---|
| Make minimal assumptions | 0.22 |
| Give option not to act | 0.21 |
| Minimize threat | 0.22 |
| Communicate want not to impinge | 0.34 |

Table 2: Frequency of autonomy oriented strategies.

result down to the nature of the power relationship between the speaker and the audience. Since both the speaker and the critics were students, the critics obviously did not feel like being in the position of judging on the performance of their colleagues.

Out of the five autonomy oriented strategies, only four can be found in the corpus (see Table 2). Apart from the communicative strategy "Communicate want not to impinge" which relies mainly on the impersonalization of the speech act (reflected by the avoidance of pronouns, such as "you" and "I") and which is used in 34% of the time, the use of the other strategies is equally distributed around 22%. Most communicative acts that correspond to the category "Make minimal assumptions" employ hedging verb phrases, such as "I think", "I guess", or "I suppose". In case of the strategy "Give option not to act", the subjunctive is widely used along with words, such as "perhaps". The strategy "Minimize threat" employs minimizing expressions, such as "a little".

Out of the approval oriented strategies only the "claim reflexivity" strategy was used regularly (47% of the time). This strategy was realized by giving reasons for the criticism and thus trying to explain to the addressee why the criticism is necessary. Although all off record strategies identified by Brown and Levinson (1987) can be found in the corpus only one is used regularly: violate manner maxim. To realize this strategy, the critics usually employed elliptical utterances.

Furthermore, we were interested in the distribution of gesture types. Out of the six gesture types that were annotated, only two are exceptional in the frequency of their use: beats and emblems (see Fig. 3). Whereas emblems can be rarely observed (3%), beats are the most frequently used gestures (26%). Emblematic gestures are self-sufficient in that they can be interpreted without any accompanying utterance. Thus, it is not astonishing to find them rarely as coverbal gestures. Beats are rhythmic gestures that emphasize words in an utterance or relate different parts of an utterance. But they might also connect different parts of an utterance thus indicating that the turn has not yet ended. Thus, the extensive use of beats might
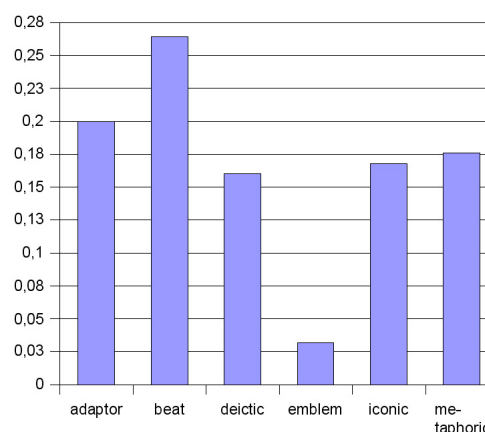


Figure 3: Distribution of gestures in the SEMMEL corpus.

be an artifact of the experimental setting because the critics had to "invent" an act of criticism that was not their own on the fly and thus the beat gesture might be an outward sign of this process indicating that the turn has not yet ended. As noted by McNeill (1992), the number of beats depends among other things on the discourse context. He observed about 25% beats in narrative contexts which roughly corresponds to our findings versus 54% beats in extra narrative contexts.

Overall, we did not notice great differences in the distribution of deictic, iconic and metaphoric gestures. However, when analyzing their co-occurrence with politeness strategies, two general tendencies may be observed (see Fig. 4). First, adaptors are used considerably while employing autonomy oriented strategies (26%). They are used least frequently with off record strategies (5%). Off record strategies are the most ambiguous and vague means to deliver a face threat. Given that adaptors often indicate that people are nervous, the more frequent use of adaptors in autonomy oriented strategies seems plausible because the criticism is delivered more openly resulting in more stress for the speaker.

Second, there is a difference in the use of gestures of the abstract (metaphoric) and gestures of the concrete (iconic and deictic). Nearly all deictic gestures that occurred in our setting referred to the addressee or concrete locations in the space (76.8%). 50% of all gestures used with the off record strategies were metaphoric in nature vs. 14% for iconic and deictic gestures. In contrast to this, 50% of the gestures employed with the direct strategies, and 49% of the gestures employed with the approval oriented strategies were iconic and deictic in nature. The same is true
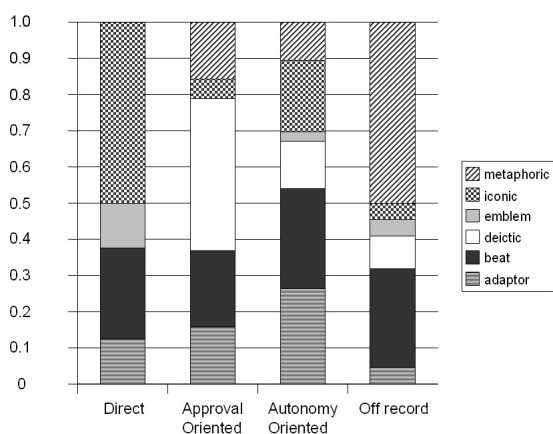
149

Figure 4: Co-occurrence of gestures and politeness strategies in the SEMMEL corpus.

to a lesser degree for the autonomy oriented strategies. In this case, 33% were gestures of the concrete and only 11% metaphoric gestures. Thus, the more abstract, vague and ambiguous the strategies become, the more abstract and vague the primarily employed gesture type becomes.

These results confirm the assumptions that not only linguistic regularities can be found in the use of politeness strategies, but that also non-verbal behaviors like gestures play a principled role in the realization of strategies. Metaphoric gestures relate to abstract concepts and illustrate an aspect of a referent in the utterance by the aid of a metaphor. The best known metaphoric gesture is the conduit metaphor where the hands form a kind of container that symbolizes the concept of a story or narrative. Most of the time, metaphoric gestures contain iconic as well as abstract parts. Why are metaphoric gestures found foremost with off record strategies? In contrast to direct strategies which do not consider the loss of face of the addressee and in contrast to approval and autonomy oriented strategies where the direct criticism is redressed but still visible, off record strategies just hint at what the speaker intends to deliver as a message, leaving the addressee at a loss to inference the speaker's intention. Being vague and ambiguous does not leave much ground for concrete gestures that refer to aspects of concrete and direct referents. Thus, metaphoric gestures are the first choice for co-verbal gestures while employing off record strategies. The contrary argument holds for the other types of strategies and the gestures of the concrete. For example, employing a direct strategy, one of the critics said: "'... some pictures of the instruments, especially of

this cornet[iconic] that you mentioned"'[1]. The direct referent cornet is iconically visualized by outlining the shape. The left hand is raised like holding the cornet, the index finger of the right hand is extended and the hand describes a circle. In the off record case the speaker might try to give only association clues, such as another critic who used an elliptical utterance: "'not so clearly to identify ... so of the structure[metaphoric] ... structure you have somehow"'[2]. Here the verbal information is accompanied by a gesture which comes in the form of the conduit metaphor. The left and right hand indicate holding something like a box.

# 4 Conclusions and future work

In this paper, we presented the results of a corpus study we devised to shed light on the question of how face threats are mitigated by multi-modal communicative acts. Unlike earlier work on politeness behaviors, we focus on how politeness is expressed by means of gestures. The results we presented are preliminary because up to now roughly half of the material has been annotated. But we are confident that the found tendencies will scale up to the whole corpus. The results indicate that gestures are indeed used to strengthen the effect of verbal acts of politeness. In particular, vagueness as a means of politeness is not only reflected by verbal utterances, but also by gestures. Iconic and deictic gestures were overwhelmingly used in more direct criticism while there was a high frequency of metaphoric gestures in off record strategies. Obviously, our subjects did not attempt at compensating for the vagueness of their speech by using more concrete gestures. Interestingly, McNeill (1992, pp. 93) noticed a high number of sequence-related iconics and deictics in narrative contexts while metaphorics appear more frequently in extra-narrative contexts. The question arises of whether the critics rather referred to the story line of the presentation in the case of direct criticism while indirect criticism rather addresses the meta narrative structure level. We will investigate this question in a further study.

The results gained from our studies may serve as guidelines for the formulation of non-verbal strategies of politeness for an ECA. We illustrate this by the BEAT system presented by Cassell et al. (2001). BEAT suggests non-verbal gestures based on a linguistic and contextual analysis of typed text. Since

---

[1]Original utterance: "'... ein paar Bilder der Instrumente, also gerade dieses Horn[iconic] dass du angesprochen hast'"

[2]Original utterance: "'nicht so klar erkennen ... so von der Struktur[metaphoric] ... Struktur habt ihr euch irgendwie'"

non-verbal behaviors are generated independently of each other, the system may end up with a set of incompatible gestures. The set of proposed gestures is therefore reduced to those gestures that are actually realized by the animation module. The findings of our studies may inform both the generation of gestures and the filtering process of the BEAT system. For instance, deictic gestures may be given a higher priority than iconic gestures when suggesting non-verbal behaviors for approval oriented strategies. On the other hand, they may be filtered out with a higher probability when realizing off record strategies. Currently, we are preparing an empirical study to compare the effect of two kinds of ECA on the user's perception of the interaction: an ECA that reflects the degree of vagueness both by speech and gestures versus an ECA that behaves inconsistently in that respect.

# References

Jens Allwood. Bodily Communication Dimensions of Expression and Content. In Björn Granström, David House, and Inger Karlsson, editors, *Multimodality in Language and Speech Systems*, pages 7–26. Kluwer Academic Publishers, Dordrecht, Boston, London, 2002.

Elisabeth Andre, Matthias Rehm, Wolfgang Minker, and Daniel Bühler. Endowing Spoken Language Dialogue Systems with Social Intelligence. In Elisabeth Andre, Laila Dybkjaer, Wolfgang Minker, and Paul Heisterkamp, editors, *Affective Dialogue Systems*, pages 178–187. Springer, Berlin, 2004.

Janet B. Bavelas, Nicole Chovil, L. Coates, and L. Roe. Gestures specialized for dialogue. *Personality and Social Psychology Bulletin*, 21:394–405, 1995.

Timothy Bickmore and Justine Cassell. Small Talk and Conversational Storytelling in Embodied Interface Agents. In *Proceedings of the AAAI Fall Symposium on Narrative Intelligence*, pages 87–92, 2000.

Penelope Brown and Stephen C. Levinson. *Politeness — Some universals in language usage*. Cambridge University Press, Cambridge, 1987.

Berardina De Carolis, Valeria Carofiglio, Massimo Bilvi, and Catherine Pelachaud. APML, a Markup Language for Believable Behavior Generation. In Z. Ruttkay and C. Pelachaud, editors, *Workshop AAMAS02: Embodied conversational agents — let's specify and evaluate them!*, 2002.

Justine Cassell, Hannes Vilhjalmsson, and Timothy Bickmore. BEAT: The Behavior Expression Animation Toolkit. In *Proceedings of SIGGRAPH '01*, pages 477–486, Los Angeles, CA, 2001.

W. Lewis Johnson, Paolo Rizza, Wauter Bosma, Sander Kole, Mattijs Ghijsen, and Hermin van Welbergen. Generating Socially Appropriate Tutorial Dialog. In Elisabeth Andre, Laila Dybkjaer, Wolfgang Minker, and Paul Heisterkamp, editors, *Affective Dialogue Systems*, pages 254–264. Springer, Berlin, 2004.

Adam Kendon. Some reasons for studying gestures. *Semiotica*, 62(1-2):3–28, 1986.

Michael Kipp. *Gesture Generation by Imitation — From Human Behavior to Computer Character Animation*. PhD thesis, Universität des Saarlandes, Saarbrücken, 2003.

Mark L. Knapp and Judith A. Hall. *Nonverbal Communication in Human Interaction*. Harcourt Brace College Publishers, Fort Worth, 4. edition, 1997.

David McNeill. *Hand and Mind — What Gestures Reveal about Thought*. The University of Chicago Press, Chicago, London, 1992.

Yukiko I. Nakano, Gabe Reinstein, Tom Stocky, and Justine Cassell. Towards a Model of Face-to-face Grounding. In *Proceedings of the Association for Computational Linguistics*, Sapporo, Japan, July 1–12 2003.

Allan Pease. *Body language — How to read others' thought by their gestures*. Sheldon Press, London, 20th edition, 1993.

Kaska Porayska-Pomsta and Chris Mellish. Modelling Politeness in Natural Language Generation. In *Proceedings of INLG*, 2004.

Helmut Prendinger and Mitsuru Ishizuka. Social Role Awareness in Animated Agents. In *Proceedings of Agents '01, Montreal, Canada*, pages 270–277, 2001.

April R. Trees and Valerie Manusov. Managing Face Concerns in Criticism — Integrating Nonverbal Behaviors as a Dimension of Politeness in Female Friendship Dyads. *Human Communication Research*, 24(4):564–583, 1998.

Marilyn A. Walker, Janet E. Cahn, and Stephen J. Whittaker. Improvising Linguistic Style: Social and Affective Bases for Agent Personality. In *Proceedings of AAMAS'97*, 1997.