Conversing with electronic devices

An integrated approach towards the generation and evaluation of nonverbal behavior in face-to-face like interface agents.

Gary Bente[1], Nicole C. Krämer[1], Georg Trogemann[2], Jens Piesk[2], & Oliver Fischer[1]
[1] University of Cologne, Department of Psychology;
http://www.uni-koeln.de/phil-fak/psych/
bente@uni-koeln.de, nicole.kraemer@uni-koeln.de, oliver.fischer@uni-koeln.de
[2] Academy of Media Arts Cologne, Laboratory for Mixed Realities
http://www.lmr.khm.de/ - trogemann@khm.de, piesk@khm.de

Keywords

Human-machine-interaction, conversational interfaces, nonverbal communication

Abstract

This paper addresses fundamental research problems concerning the implementation of nonverbal communication channels in so-called conversational interfaces: As a solution to these problems a new methodology is introduced that makes possible the generation, systematic variation and experimental evaluation of nonverbally interacting interface agents in different task contexts. A taxonomy of research designs is presented and first applications of the new methodology are summarized and discussed.

1 Introduction

The permeation of information technology through practically all areas of both professional and private lives has brought with it the challenge to develop interfaces that make the use of computers and other kinds of electronic devices easy and efficient, not only for experts but for everyone. Although pointing devices, menus and icons enable us – at least to some extend - to interact intuitively with a computer, there is still a certain expertise needed to use these input devices efficiently. Moreover, there are other restrictions, both internal (e. g. a lack of technical experience or learning motivation) and external (e. g. complexity of the task and other specific situational limitations) that can make an electronic device a hard system to handle even when a menu driven interface is available (imagine a 85 year-old man trying to program his satellite receiver). Aware of these problems, recent interface developments try to incorporate features of human face-to-face-communication into human-machine-interaction to increase flexibility, efficiency and acceptance. Especially life-like interface agents, using natural speech and nonverbal behavior (gestures, facial displays, etc.) are considered to be a promising approach for modern interface design [1,2,3]. These developments are based on the assumption that interfaces which reproduce the functionality of face-to-face-interaction are more universal and allow an intuitive approach even to new and complex systems. Although this assumption is very plausible at first sight, the empirical evidence for this hypothesis is still lacking.

We might find some anecdotical evidence for the usability of a verbal input device, e.g. for a voice control in a car radio. But when it comes to the integration of nonverbal phenomena (gestures, smiles and head nods) into conversational interfaces, there are hardly any such examples hinting to a better usability of the system. Moreover, seeming advantages of specific nonverbal input and output options, such as pointing gestures, gaze direction or illustrative movements, might be neutralized or even predominated by potentially negative side effects of the visual appearance of an embodied system agent: For example the user might feel distracted from the work in progress, or just dislike the agent's appearance, or the way it behaves. Indeed, recent studies highlight the psychological impact of embodied interfaces, but they also point to certain risks. For example Sproull et al. [4] could show that users attributed personality traits to human-like interfaces, were more aroused in the presence of animated characters, and tried to present themselves in a more positive light.

Rickenberg and Reeves [5] similarly found a higher "trustworthiness" attributed to such a system, but they also observed more self-monitoring, a higher level of anxiety, and more mistakes in the solution of the task. Their results also indicate that these effects depend on the personality traits of the users.

Against this background Parise et al. [6] point out the importance of monitoring specific psychological effects in the development of face-to-face-like interfaces: *"As computer interfaces can display more life-like qualities such as speech output and personable characters or agents, it becomes important to understand and assess user's interaction behavior within a social interaction framework rather than only a narrower machine interaction one" (p. 123).* Within this framework it then becomes evident that the specific dynamics of the nonverbal behavior deserve more attention than the mere presence or absence of the 'Virtual Vis-a-Vis' (3V). Rickenberg and Reeves [5] conclude: *"... that decisions concerning the use of animated characters should address the details of execution and social presentation"* and that it is not sufficient *"... to focus on whether or not an animated character is present. Rather the ultimate evaluation is similar to those for real people - it depends on what the character does, what it says, and how it presents itself (p. 55)".*

The scientific observation and analysis of the different behavioral variations and their specific effects, however, has to allow for a wide range of human and situational factors. So the question is: "Which specific interface qualities lead under which conditions, in which people, to which effects?" [5]. To answer this question a methodology is needed that does not only allow the generation and experimental variation of static and dynamic features of the virtual actors (visual appearance and nonverbal behavior), but also the measurement of different dependent variables, such as performance, emotional arousal, communication effort, ease of use, liking, etc. Within the interdisciplinary research program EMBASSI, which is funded by the German Ministry of Education and Research (BMBF), expert groups representing the disciplines psychology, computer science, and media-art have started a joint development of a universal platform for the computer generation of nonverbal behavior. It is set within the framework of conversational interfaces and thus provides an adequate basis for a systematic evaluation of such an interface in different application areas as well as with different user groups. In the present paper, we will report on the objectives of the EMBASSI project and the first results obtained from evaluation studies.

2 Functional aspects of NVB in embodied interface agents

Nonverbal behavior (NVB) represents a very subtle and extremely complex communication system in human interaction which can hardly be modeled in analogy to the principles of speech production and understanding. Many signal subsystems, such as facial activity, body posture, head movements, gestures, etc. occur simultaneously in FTF-interaction and can serve different interpersonal functions at the same time. These communicative functions should be at least roughly differentiated, as this is a crucial basis both for a computer based modeling approach and for a later analysis of possible effects and side effects when implementing nonverbal behavior in embodied interfaces. Table 1 points out four functional aspects of nonverbal behavior in FTF interactions that are likely to be relevant for development of animated interface agents as well. It also depicts the names of researchers who have worked on these aspects either within FTF-interaction or who have indeed implemented these functions within a Human Computer Interaction (HCI). The classification of the different functions is crucial to an understanding of the psychological relevance of different types of behavior and will hence be described in more detail.

Table 1: Functional aspects of nonverbal behavior in face-to-face and human-computer interaction

Functions	FTF-interaction	HC-interaction
Modeling functions	Bandura [7]	
Discourse functions	Bolinger [8] McNeill [9] Chovil [10]	Cassell et al. [11], Nagao & Takeuchi [2]
Dialog functions	Duncan [12]	Cassell et al. [1] Thórisson [3]
Relational functions	Mehrabian [13] Exline et al. [14] Frey [15]	

Modeling functions are connected to the fact that humans seem to have clear advantages in performing motor tasks when they can observe somebody else showing the required movements [7]. This function has been discovered in the field of character animation some time ago (see for example computer animations illustrating safety instructions in airplanes). This function deserves attention in the context of anthropomorphic interfaces when specific motor skills are the focus of the requested information (e.g. when instructing the user how to handle an electronic device in a private household).

Discourse functions are closely related to verbal behavior and can work either as complements, supplements or substitutes of speech. Pointing gestures, illustrative gestures, beat gestures (such as waving a hand to structure the speech flow) or emblematic gestures (such as symbolizing a certain object) belong to this functional category. Within 3V-interfaces nonverbal behaviors like these could possibly increase understanding and help to structure the processing of verbal information.

Dialogue functions include turn-taking signals (e.g. eye contact) and back-channel signals (e.g. head nods) and serve the smooth flow of interaction when exchanging speaker and listener roles. In human computer interaction these signals could be important on the input as well as on the output side, e.g. when the user should understand that a verbal input is needed or when interrupts and feedback loops must be generated because the user has given negative back-channel signals to indicate that something is still not understood.

Relational or socio-emotional functions are the least explored in nonverbal communication research. Although ample empirical evidence proves that nonverbal behavior has an enormous impact on our perception and evaluation of other people (how much we like them, whether they are trustworthy or not, etc. [16]) little is known about the secret codes that actually lead to those interpersonal effects. Just as in FTF-interaction, nonverbal behavior in animated characters could induce positive feelings in the human user, increase motivation and thus facilitate task performance [17,18]. But, as in every day life, the behavior of the computer actor can also do exactly the opposite: evoke negative feelings and impair task performance. Given these risks and the restricted state of knowledge in this area one might suggest to skip this functional aspect and to focus the discourse and dialog functions. It can be seen from table 1 that this is exactly what most researchers have decided to do.

However, what has been neglected is the fact that the visual presence of the computer actor itself will necessarily lead to a socio-emotional reaction of the human vis-a-vis, whether or not these effects were intended by specific behavioral variations. This means that a lack of gestures or facial activity will *not* be attributed to a technical restriction of the system, but to the specific personal characteristics of the virtual vis-a-vis. This means that it will rather be perceived, for example, as cold or aloof, which could lead to the avoidance of any further interaction with the agent. Exploring all theoretical aspects of the socio-emotional functions is way beyond the scope of this article. However, we want to stress that this interpersonal dimension has to be taken into account from the very beginning of the development process. If we can't use it to our advantage, we should at least avoid the negative side effects wherever possible.

3 Top-down vs. bottom-up modeling of non-verbal behavior (NVB)

Extending the scope of research to socio-emotional dimensions does not only put difficult theoretical questions, but it also reveals specific limitations of current methodology. Traditional approaches to the implementation of NVB in human computer interaction as listed in table 1 are based on so called *top-down-modeling-* or *theory driven* approaches. This means that they start from already well investigated aspects of human behavior and convert these into clearly formulated, mathematical rules, telling which behavior should occur in which phase of an interaction with which particular effects. In the area of dialogue or discourse functions, there is at least a minimal body of knowledge on which algorithms for behavior generation can be based. For the area of socio-emotional functions such knowledge is virtually non-existent. Consequently, top-down modeling approaches are very unlikely to issue embodied computer agents that are perceived as being face-to-face-like, simply because their nonverbal behavior is very restricted. Top-down approaches thus might be functional in testing algorithms for the generation of a limited set of nonverbal cues in isolated experimental conditions, but in practical applications the negative socio-emotional side effects will most probably make them rather useless.

These restrictions do not occur within *bottom-up modeling-* or *data driven* approaches. Bottom- up approaches are based on detailed transcripts of real-life interactions and thus allow the computer generation of natural nonverbal interaction behavior, which evokes reactions in the user that are as close to the real situation as possible [19]. However, there is a downside to this approach as well: Nonverbal behavior cannot be generated autonomously, it has to be retrieved from a prerecorded database. To make this 'canned behavior' accessible in a human-machine interaction, specific annotations have to be added to the database that can be used as an index to the descriptive behavior data. These annotations can serve as a reference to a complex motor task, to a speech context, to an interpersonal situation in which the nonverbal behavior should occur, or to an intended emotional response. With respect to the socio-emotional effects of NVB, the annotation problem can be addressed empirically, e.g. by asking neutral observers for their impressions of a protocol-based computer animation. In such experiments systematic variations of specific nonverbal cues can be incorporated to test their particular effects on person perception and impression formation.

Thus, bottom-up- and top-down-approaches do not necessarily exclude each other. The inductive research strategy that is characteristic for the bottom-up approach can lead to new insights in the functional principles of NVB, which can in turn be used to formulate algorithms for its autonomous generation within a subsequent top-down approach. Aiming at the integration of both strategies, we have developed a new methodology for the computer generation and experimental evaluation of NVB in embodied interface agents. The result of our effort, the 'Development and Evaluation Platform for Animated Characters (DEPAC)' will be described in the following section.

4 The Development and Evaluation Platform for Animated Characters (DEPAC)

DEPAC provides the necessary tools for recording, editing, and displaying nonverbal behavior, as well as tools for annotation, experimental variation and evaluation. It aims at establishing an exhaustive behavior database (motion base) that supports the development of interactive virtual actors and the process of evaluation experiments. The combination of top-down and bottom-up approach, as described above, should then enable the implementation of psychological and linguistic rules for autonomous computer simulation of life-like nonverbal behavior in HCI (rule base). DEPAC is structured in 5 layers: Input layer (I), Editor layer (E), Output layer (O), Filter & Generator layer (F&G rule base), Dialogue layer (D), Record&disPlay section (R&P motion base) and a cross-module utility section (US).

(I) Input layer: The input layer allows the detection of movement behavior by either using automatic measurement devices such as magnetic and optical motion-capture systems, data-gloves or eye-tracking systems or by means of video transcription. Data are stored within the motion base in Biovision-Format (.bvh-Files). When audio-data and facial activities are recorded simultaneously a proprietary xml-Format is used. The input layer incorporates a basic audio recording section which is synchronized with the motion recording.

(E) Editor layer: A tool-set for cutting, editing and annotating human motion and voice is provided by the editor layer. The editor layer provides direct access to the motion and the rule base. Different editors such as numerical ASCII-editors, function curve editors and graphical editors using inverse kinematics are part of this layer. The editor layer encompasses an annotation section that can be accessed by the dialogue layer (see figure 1).

Figure 1: Editor layer with real-time 3D-player

(O) Output layer: A real time player for interactive display of recorded or synthesized speech and nonverbal behavior is integrated in the output player. The output layer encompasses tools for mapping norm skeleton motions to other character skeletons as well as tools for choosing between different rendering and lighting options and texture mapping. Blending functions are included in this layer as well.

(F&G rule base) Filter/Generator layer: Rules for behavior construction and/or selection and a taxonomy of situations and tasks with varying performance and socio-emotional qualities are stored within the rule base and accessed for editing purposes by the editor layer. During generation of experimental settings or during interaction the rule base interfaces between output layer and motion base.

(D) Dialogue layer: The dialogue layer provides tools for connecting motion and rule base with input and output layer. It uses the annotation section to select special motion sequences and to integrate them into the dialogue. Approaches to dialogue management can be broadly classified into finite-state methods, on the one hand, and self-organizing, inference-based approaches on the other [20, 21]. DEPAC concentrates first on scripted dialogues with finite-state methods and will develop on later stages towards free style dialogues with self-organizing approaches.

(R&P motion base): The record and disPlay section contains the motion base with prerecorded behavioral data for body movement, facial activity and lip movements as well as the semantic annotations of the stored data. The annotations serve as indexes to the behavior batches that can be accessed by he dialogue manager.

(US) Utility section: The main focus of the utility section is interfacing between other commercial and non-commercial modeling, animation and rendering engines and standards. The utility section also provides instruments for real time measurement of physiological arousal as well as a battery of rating scales and questionaires for evaluation.

The workflow concept of building motion and behavior libraries is shown in Figure 2: After storing the recorded sequences within an experimental database the functionality of the editor layer (E) is used to build specific variations of movements according to the experimental setting. These variations are presented by means of the output layer (O) to the participants of the impression experiment. The outcome of this evaluation process is a set of evident motion impressions which can be selected, cut and annotated by the editor layer (E) for storage within a motion or behavior library.

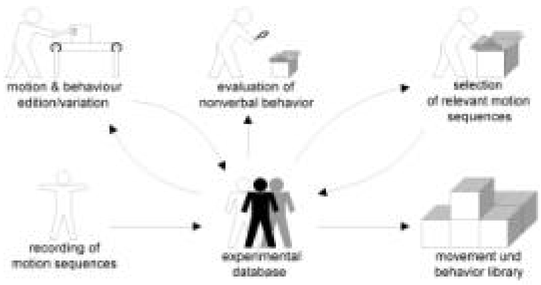

Figure 2: Workflow concept of DEPAC

5 Research paradigms for usability tests and evaluation

DEPAC is constructed as an open system that can be used for different evaluation tasks on different stages of development. The application of the various research designs will depend on two factors: (1) the knowledge on nonverbal communication and dialog management and (2) the availability of advanced input and output devices. These include e.g. video-based recognition systems for gestures and facial activities and facilities for the real time generation and synchronization of verbal and nonverbal behavior. Four design types can be distinguished at the movement:

(1) *Third Party Observation (TPO):* Participants are in a passive observation situation. Interactions between two computer generated characters are presented introducing specific variations in the NVB of the virtual actors. There is no interactivity between user and virtual actors. The design is used to generate basic knowledge on the interpersonal effects of specific nonverbal cues. An example of this design type is given in section 6.

(2) Hidden Expert Dialogs (HED): The user is interacting with an expert next door, who is connected to motion capture devices and represented by a real time 3D model to the user. User actions are transmitted to the expert via audio/video connection. This setup guarantees full interactivity and thus allows to explore the effects of anthropomorphic interfaces under optimal conditions. Appearance of the animated figure can be varied experimentally.

(3) Script Driven Interaction (SDI): The user is polling prerecorded behavior from the data base. The design allows for semi-interactivity. The interface agent responds to certain requests of the user (e.g. asking for information on how to program a video recorder). Complete sequences of behavior then are played back from the data base. Systematic variations of static (appearance) and dynamic (nonverbal behavior) cues can be applied. User requests in the first step are send to the system via mouse click. Next generation SDIs will allow to use natural speech and gestures also on the input side.

(4) Free style conversations (FSC): Full functionality of DEPAC will include a dialogue management system based on semantic protocols of user speech and nonverbal behavior and thus allow for the experimental analysis of fully interactive systems. The realization of such a system will depend on the technological advancements in the area of speech and gesture recognition as well as on growing knowledge about the principles of spontaneous production of speech and gestures in face-to-face dialogs.

6 First research applications of DEPAC

TPO was the first of the four settings to be realized in DEPAC. In a first experiment the approach was tested by analyzing the effects of three specific nonverbal cues: head tilt, general activity of head movement, and trunk position. Three one minute dyadic interactions were transcribed from video recordings according to the principles of the Bernese Coding System [22]. The three sequences differed strongly in terms of both global activity of the actors and socio-emotional climate: While the first sequence showed a rather emotional, conflict laden interaction between supervisor and employee, the second interaction could be characterized as a casual chat, with the left stimulus person being rather passive and mostly in the role of the listener. In the third interaction, also a casual chat, the left person was much more active while the interaction partner stayed in the listener's role for most of the time.

Experimental variations

Systematic variations of head posture, general movement activity, and trunk position were inserted into the data protocols while all other aspects of behavior were held constant. The experimental variations concerned only one of the interacting figures (the left person on the screen). We decided to experimentally vary the head tilt, because it has repeatedly proven to be a very powerful cue in person perception: It influences both the perceived socio-emotional distance and the perceived power of the person [22]. Our experiment included three different variations of the head tilt, with the first being the ‘normal’, upright position (see figure 3). For the second variation, the head was tilted 20 degrees towards the interaction partner (see figure 4), and for the third variation it was tilted 20 degrees away (see figure 5). It is important that we only changed the mean position of the head while the dynamic head movement itself was left unchanged.

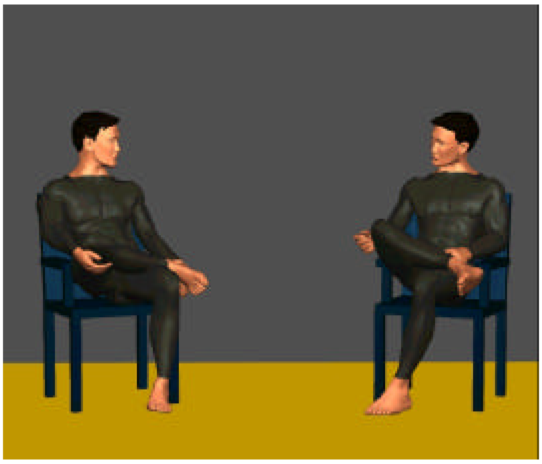

Figure 3: The Person on the left with the head *upright*

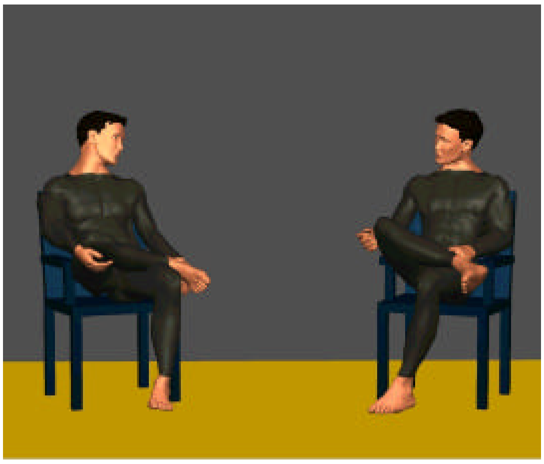

Figure 4: The Person on the left with the head tilted *towards* the interaction partner

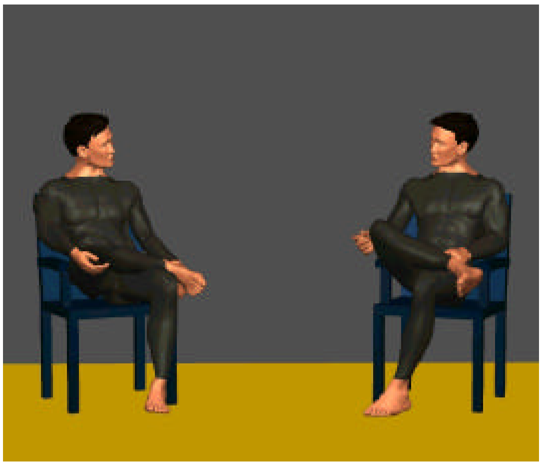

Figure 5: The person on the left with head tilted *away* from the interaction partner

In addition to the head position, we varied the general level of head movement activity, which is another important although rather unexplored cue in interpersonal communication [23]. This was done by extending the end points of all movements, so that the movements were both intensified and accelerated. The last experimental condition concerned the leaning of the trunk: It was put to a constantly upright position. Specifically, in two sequences the trunk leaning was changed from backward to upright (see e. g. figure 6), and in one sequence the leaning was changed from forward to upright. The trunk position is not only a so-called immediacy cue [13] and thus strongly related to evaluative judgements, but it is also a dominance cue, meaning that it has an influence on how powerful the person is perceived.

Figure 6: Person on the left with the trunk *upright*

Participants and Procedure

160 students of the University of Cologne, differing in age and sex participated in the study. An independent group design was chosen, presenting the three animation sequences to each group. The different experimental variations were systematically assigned, making sure that each group saw a different combination of the experimental variations. All sequences were presented without sound. After each interaction sequence, participants were asked to evaluate both actors on the screen. This was done by means of a 36 item bipolar adjective scale, containing answers on how friendly, intelligent, dominant, strong, nervous, active etc. the person was perceived. The items were selected in a pilot study and represented the three basic dimensions

in person perception: *evaluation, dominance and activity* [13].

Results

To illustrate the impact of even slight variations in nonverbal behavior only some of the significant results are reported here. When the stimulus person has tilted the head away from the interaction partner he is evaluated more positively than with the head tilted towards the interaction partner. This can be shown for all three sequences. In sequence 1 the effect occurs for the item "not likeable-likeable" (ANOVA: F = 4,54; p = ,002). As can be seen from figure 7, the stimulus person is perceived as more likeable when the head is tilted away from the interaction partner than when it is tilted towards the partner (post hoc Scheffé: p = ,022) or more active (p = ,038).

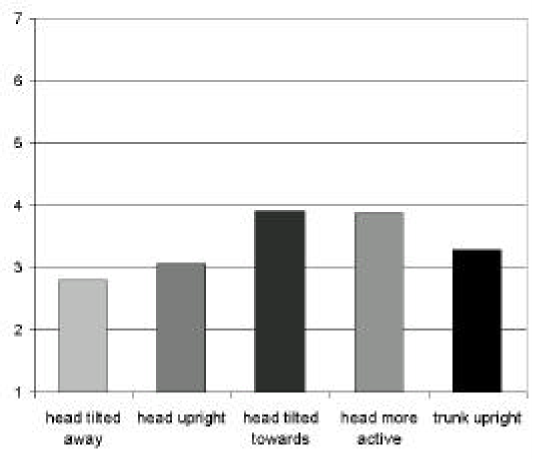

Figure 7: Mean judgements for the stimulus person in sequence 1 concerning the item "not likeable"

In sequence 3 a significant difference can be found for the item "indifferent-sympathetic" (F = 3,17; p = ,015): again the stimulus person was perceived as more positive when head is tilted away from the interaction partner than when it is tilted towards the partner (p = ,029).

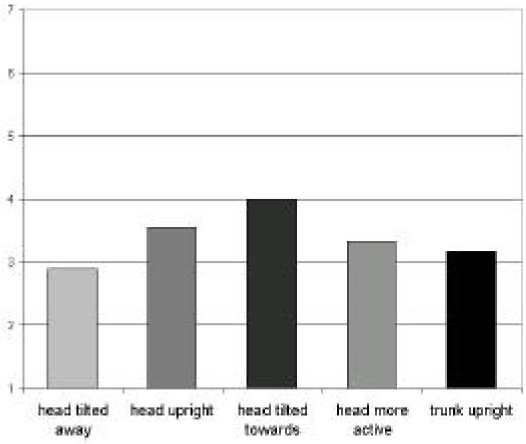

Figure 8: Mean judgements for the stimulus person in sequence 3 concerning the item "indifferent"

In sequence 1 this phenomenon is apparently related to the fact that the person with the head tilted away from the interaction partner is perceived as being more dominant than the one with the head tilted towards him. Investigation of the 'dominance' items shows that the rating of dominance is linearly related to the degree of head tilt (see figure 9). Concerning the item "weak-strong" (F = 6,6; p = ,000) the person is rated as less strong in the conditions "head tilted towards" (p = ,000) and "head upright" (p = ,041) than in the condition "head tilted away". The correlation between positive evaluation and dominance is puzzling, but only at first sight: As the stimulus person in this case is generally perceived as being inferior to the interaction partner, its dominance signals make the situation more equal and thus less unpleasant to observe.

Effects of posture are both consistent for all sequences and in accordance with earlier findings [24]. When stimulus persons are sitting upright instead of leaning backwards they are perceived as less strong, dominant and relaxed, but as more involved and attentive.

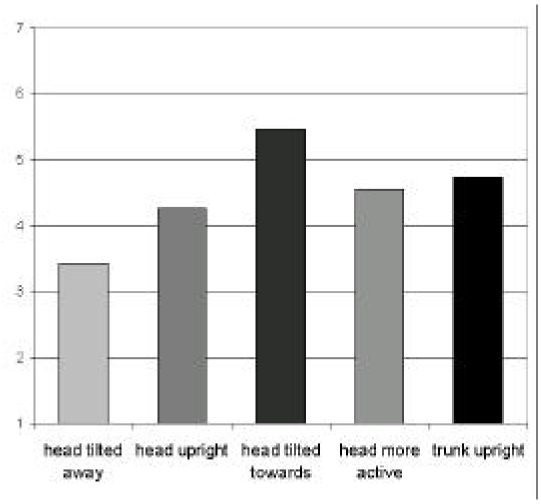

Figure 9: Mean judgements for the stimulus person in sequence 1 concerning the item "weak"

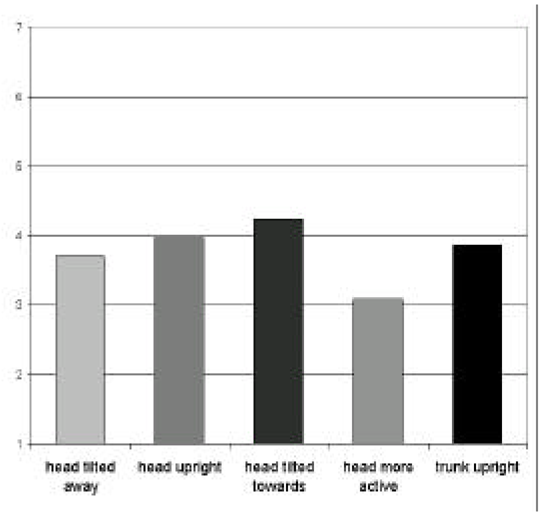

Figure 10: Mean judgements for the stimulus person in sequence 2 concerning the item "not likeable"

Variations of head movement intensity did not produce consistent results. While in sequence 1 the person is perceived as less likeable when showing more pronounced head movements (see figure 7), the stimulus persons in sequence 2 and 3 (with equally increased head movement) are evaluated more positively. For example, in sequence 2 (see figure 10) the person is perceived as more "likeable" ($F = 2{,}74$; $p = {,}031$). These inconsistent effects can obviously be attributed to differences in the social situation: Whereas in sequence 1 the situation was apparently conflict laden and aggressive, the interactions in sequence 2 and 3 appear to be rather friendly chats. This makes a clear point for considering a strong context dependency when using semantic annotation for nonverbal behavior in our data base.

7 Conclusion and perspective

The results from our pilot study point to a strong impact of even subtle nonverbal phenomena on the perception and evaluation of virtual characters. However, as the data presented here are based on a 'Third Party Observation (TPO)' design, it has yet to be verified that the effects obtained for head tilts and body leaning do also occur when the virtual vis-a-vis is directly addressing the user. Such experiments should now be conducted as a second step within 'Script Driven Interaction (SDI)' designs, where the obtained results can be used to insert nonverbal variations into different application contexts. First SDI-studies are conceptualized as short interactions between a user and a personal video recorder assistant. This setup will not only lead to a higher degree of interactivity but will also make it possible to systematically explore the influence of context and person variables (such as task difficulty, expertise, arousal etc.) on the efficiency, effectivity and acceptance of embodied interface agents.

References

[1] Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K, Vilhjálmsson, H., Yan, H. (1999). Embodiment in conversational interfaces: Rea. CHI´99 Conference Proceedings (pp.520-527). Association for Computing Machinery.

[2] Nagao, K. & Takeuchi, A. (1994). Social Interaction: Multimodal Conversation with Social Agents. 12th National Conference on Artificial Intelligence (AAAI), 22-28.

[3] Thórisson, K. R. (1996). Communicative humanoids. A computational model of psychosocial dialogue skills. PHD-Thesis, MIT.

[4] Sproull, L., Subramani, M., Kiesler, S., Walker, J.H., & Waters, K. (1996). When the interface is a face. Human Computer Interaction, 11 (2), 97-124.

[5] Rickenberg, R. & Reeves, B. (2000). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. Letters of CHI 2000, April 2000, 49-56.

[6] Parise, S., Kiesler, S., Sproull, L., & Waters, K. (1999). Cooperating with life-like interfaces. Computers in Human Behavior, 15, 123-142.

[7] Bandura, A. (1977). Social learning theory. Englewood Cliffs: Prentice Hall.

[8] Bolinger, D. (1983). Intonation and gesture. American Speech, 58 (2), 156-174.

[9] McNeill, D. (1992). Hand and mind: what gestures reveal about thought. Chicago: University of Chicago.

[10]Chovil, N. (1991). Discourse oriented facial displays in conversation. Research on Language and Social Interaction, 25, 163-194.

[11]Cassell, J., Steedman, M., Badler, N. Pelachaud, C., Stone, M. Douville, B., Prevost, S. & Achorn, B. (1994). Modeling the interaction between speech and gesture. In A. Ram & K. Eiselt (eds.), Proceedings on the sixteenth anuual conference of the cognitive science (pp. 153-158). LEA.

[12]Duncan, S., Jr. (1972). Some signals and rules for taking speaking turns in conversations. Journal of Personality and Social Psychology, 23 (2), 283-292.

[13]Mehrabian, A. (1970). A semantic space for nonverbal behavior. Journal of Consulting and Clinical Psychology. 35 (2), 248-257.

[14]Exline, R. V., Ellyson, S. L. & Long, B. (1975). Visual behavior as an aspect of power role relationships. In P. Pliner, L. Krames & T. Alloway (Eds.), Nonverbal communication of aggression (pp. 21-52). New York: Plenum.

[15]Frey. S. (1999). Die Macht des Bildes. Göttingen: Huber.

[16]Mehrabian, A. & Ferris, S. R. (1967). Inference of attitudes from non-verbal communication in two channels. Journal of Consulting Psychology, 31, 248-252.

[17]Lester, J. C., Converse, S. A., Kahler, S.E., Barlow, S. T., Stone, B. A. & Bhogal, R. S. (1997). The persona effect: affective impact of animated pedagogical agents. In S. Pemberton (Ed.), Human Factors in Computing Systems: CHI´97 Conference Proceedings (pp. 359-366). New York: ACM Press.

[18]Lester, J. C., Converse, S. A., Stone, B. A., Kahler, S. E., Barlow, S. T. (1997). Animated pedagocical agents and problem-solving effectiveness: A large-scale empirical evaluation. In B. DuBoulay & Mizoguchi (Eds.), Proceedings of the 8th World Conference on Artificial Intelligence in Education (pp. 23-30).

[19]Bente, G., Petersen, A., Krämer, N.C. & Buschmann, J.-U. (1999). Virtuelle Realität im Forschungseinsatz. Ein Wirkungsvergleich videovermittelter und computersimulierter nonverbaler Kommunikation. Medienpsychologie: Zeitschrift für Individual- und Massenkommunikation, 2, 95-120.

[20]Aretoulaki, M., Ludwig, B. (1999): Automaton-Descriptions and Theorem-Proving: A Marriage made in Heaven?, Proceedings of the IJCAI-99 Workshop on Knowledge and Reasoning in practical dialogue systems, Stockholm, Sweden, 9-16.

[21]McTear, M. (1998) Modelling Spoken Dialogues with State Transition Diagrams: Experiences with the CSLU Toolkit. Proc 5th International Conference on Spoken Language Processing, Dec 1998, Sydney, Australia.

[22]Frey, S., Hirsbrunner, H. P., Florin, A., Daw, W. & Crawford, R. (1983). A unified approach to the investigation of nonverbal and verbal behavior in communication ressearch. In: W. Doise & S. Moscovici (Eds.). Current issues in European Social Psychology (S. 143-199). Cambridge: Cambridge University Press.

[23]Grammer, K., Fieder, M, & Filova, V. (1996). The communication paradox and a possible solution: Toward a radical empiricism. In A. Schmidt, K. Atzwanger, K. Schäfer & K. Grammer (eds.), Advances in human ethology, (pp. 91-120). London: Plenum.

[24]Mehrabian, A. & Williams, M. (1969). Nonverbal concomitants of perceived and intended persuasiveness. Journal of Personality and Social Psychology, 13 (1), 37-58.