Thilo Paul-Stueve (Hrsg.): Mensch & Computer 2007 Workshopband. Weimar: Verlag der Bauhaus-Universität Weimar. 2007, S. 25

„Helen“ – Embodiment in Automobile Speech User Interfaces

Matthias Görtz[1], Thomas Mandl[1], Luis Arévalo[2], Christa Womser-Hacker[1]

[1]Information Science, University of Hildesheim, Germany; [2]Blaupunkt GmbH

Abstract

Embodied conversational agents may add value to speech dialogues within driver information systems. The experiment reported in this paper, compares a speech controlled driver information system developed by Blaupunkt GmbH with a prototype enhanced with a virtual character provided by Charamel GmbH. An expert evaluation and a usability study were conducted in order to determine whether an anthropomorphic interface leads to a more effective dialogue and to a higher level of user acceptance.

1 Introduction

The ultimate goal of user-centred interface design is to achieve intuitive user interaction. The various guidelines that have been proclaimed in order to guarantee a user-friendly interface lead to the question whether or not the human-machine interaction should be simulating the most natural way of communication to the user, interpersonal interaction. Shneiderman's (1998) guidelines emphasize the difference between human-machine and human-human interaction. „Avoid presenting computers as people“ (Shneiderman 1998, 385) is one of his major claims. Latest developments in the interface design nevertheless deal with the human-like representation of computer systems (Cassell 2001, Krämer & Bente 2000, Rist et al. 2004). They consider „embodied conversational agents as a promising option for presenting information to users“ (Rist et al. 2004, 377). Especially the use of multiple modalities and the simulation of a familiar conversational situation are considered *noise-reducing* factors that lead to better usability. This paper contributes to this debate. A comparative usability study is presented which examined the user's interactive behaviour and performance with an anthropomorphic interface. Additionally, the user's acceptance was challenged by a very sensitive environment: the automobile. A prototypical modification of an existing speech controlled driver information system was enhanced by a virtual character, called „Helen“.

2 „Helen“ – The Embodied Conversational Agent

Within the Automotive environment voice user interfaces have become an important technology to manipulate in-car applications, such as the radio, navigation system, cd-player or mobile telephony. The main focus of these interfaces is to enable „hands-free“ control of so-called „comfort features“, that are not vital to the manipulation of the driver’s main task: driving the car. In this context, voice controlled systems are broadly established and throughout past years evolved from single-word commands to more mature speech dialogue systems. Reaching for the ultimate goal of natural speech interfaces, these systems shift the interactional paradigm from human-machine to simulated human-human. „Helen“ is the name of an embodied conversational agent actually taking this last step to overcome the traditional paradigm of man-machine interaction by presenting a driver information system’s features with the human-like character of a female information agent.

2.1 Visualization

„An embodied conversational agent [is] an interface in which the system is represented as a person, information is conveyed to human users by multiple modalities such as voice and hand gestures“ (Cassell 2001, 67). Another term within this field is the anthropomorphic interface agent, which highlights the aspect of human-like appearance.

Figure 1: Embodied Conversational Agent during destination input dialogue

Figure 1 displays the character-enhanced prototype used within this study. The human-like body language is used to fulfill certain functions any visual interface contains in order to reveal the system’s status or highlight certain areas of the screen. The screenshot shows two steps within the navigation system. The user needs to select and confirm the street name entered before.

2.2 Hypotheses

As mentioned in the introduction, within usability research there is an analytical discourse on whether the described paradigm shift of simulating interpersonal communication indeed leads to better usability in terms of user performance and acceptance. As a consequence, this

study adopts Cassell’s statement as a basis for its hypotheses. „… we propose to leverage the full breadth and power of human conversational competency by imbuing the computer with all of the conversational skills that humans have; to whit, the ability to use the face, hands, and melody of voice to regulate the process of conversation“ (Cassell 1999: 520) Following this argumentation the study assumed the following outcomes:

- Using a human-like representation of the driver information system within the speech dialogue will *simulate clues of interpersonal communication* and thus *lead to an intuitive usability* for first-time users.
- The elaborate *and high quality visualization* of the interface increases customer attraction and also *enhances user acceptance* of voice controlled driver information systems across technology-affine user groups.

3 Testing User Performance and Acceptance of „Helen“ in the Automotive Environment

3.1 Evaluation Design

In order to prove or disprove the above mentioned hypotheses a two-staged evaluation design was conducted. First, expert interviews were conducted. This heuristic evaluation analyzed the two prototypes of Blaupunkt GmbH and Charamel GmbH with the objective of identifying major technical differences and possible influence factors on the user’s acceptance respectively performance. Based on the experts’ findings, the usability lab set-up was designed and the prototypes as well as the work hypotheses refined. In a second step, 31 subjects across defined demographic parameters took part in a simulated interaction with both prototypes. The users had to complete certain tasks with a voice user interface displayed only on a schematic interface, and other tasks with the same degree of complexity with the same speech engine, but different visualization in the shape of „Helen“. During these lab exercises the subjects were monitored by camera, one test supervisor, and measuring devices recording the user’s electrodermal activity and heart rate. The main figures monitored task performance data, such as task completion time, task error rate, task errors, and recognition rate. In addition to the quantitative figures subjective measures were surveyed in post-task questionnaires evaluating the user experience during the tasks and acceptance of the different prototypes.

3.2 Key Findings

Across different user groups, the overall subjective response to the human-like interface was quite negative. For example, 69% per cent of the subjects with a low degree of technical experience strongly rejected the prototype with the virtual character. „The lady, floundering around, is irritating me. She neither helps nor is she nice to look at or moves her lips simultaneously to the speech output.“ (Subject statement). This typical statement from the post-task questionnaire illustrates the high degree of rejection of the animated persona on the

screen of the driver information system. The negative response to the „Helen"-prototype might implicate that the users are disappointed about the system's performance and ease of use. The performance data, however, showed a significant positive effect of the embodied agent and measurable differences in the effectiveness of use across subject characteristics, such as age, gender, and technical know-how.

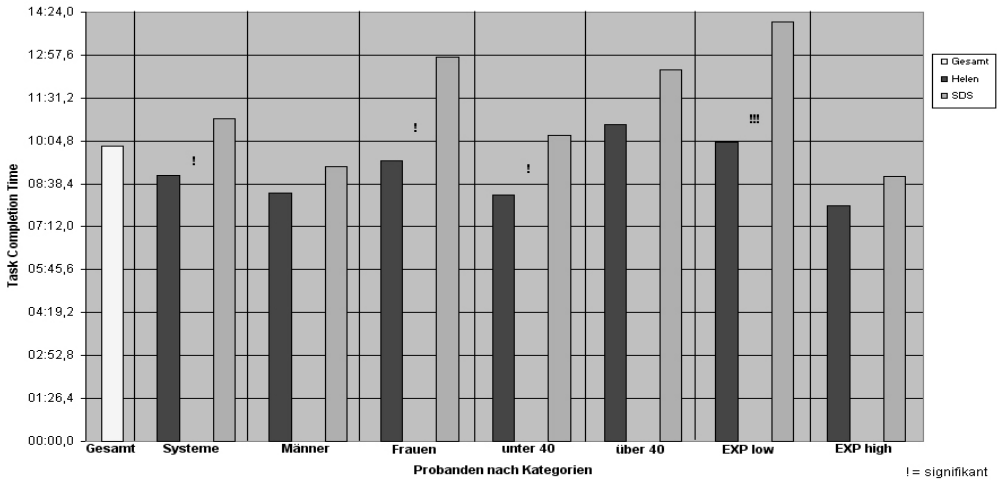

Figure 2: Task Completion Time shows significant performance increase with „Helen"

Figure 2 shows as an example the task completion time comparing the speech dialogue system (SDS) and the same technological prototype enhanced with the virtual character of „Helen". The average task completion time dropped from 10:49 minutes with the schematic user interface to 8:54 minutes with the „Helen"-prototype. During the interaction with the embodied conversational agent, the subjects made less input errors, had a higher task completion rate, and as shown in the graph completed the given tasks in significantly less time. Especially the female user group, users under 40 years of age, and users with a low degree of technical experience interacted actually more efficiently with the driver information system, when the system was augmented with the embodied conversational character.

4 Conclusion

Even though the subjective rating of the prototypes expresses a strong rejection of such an animated character in the driver information system, user performance data and therefore an essential element of system usability points in the opposite direction. The position of Cassel (1999) highlighting the positive effects of interactional clues and body language of an embodied conversational agent simulating the interpersonal communication could be proven. Nevertheless, the key to a system's usability is its user acceptance. Especially the automotive environment underlies strong restrictions. The eSafety-initiative of the European Commission postulates in its European Statement of Principles on the Design of Human Machine

Interaction that visual animations within the driver's field of vision constitute a possible distraction and therefore endanger the driver's safety. These concerns seem to influence user awareness to a large extent and constitute a sensitive barrier for integrating embodied conversational agents in driver information systems. However, the results of this study provide promising hints concerning the use of embodied conversational agents and task performance in speech dialogue systems. Further investigation in the interplay of user performance with user acceptance and its measurability will be required to deepen the understanding of usability in a certain applicational context (e.g. *automotive*) and lead to conclusions on possible implementation scenario of such embodied interfaces as „Helen“.

Literaturverzeichnis

Cassell, J.; Thórisson, K. R. (1999): The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. In: Journal of Applied Artificial Intelligence, 13 (3). Taylor & Francis, pp. 519-538.

Cassell, J. (2001): Embodied Conversational Agents. Representation and Intelligence in User Interfaces. In: AI Magazine. Winter 2001. Vol 22 (4). Menlo Park, CA: AAAI, pp. 67-84.

Krämer, N.C.; Bente, G. (2002): Virtuelle Helfer: Embodied Conversational Agents in der Mensch-Computer-Interaktion. In Bente, G.; Krämer, N.C.; Petersen, A. (Eds.) (2002): Virtuelle Realitäten. Göttingen: Hogrefe Verlag für Psychologie, pp. 203-225.

Rist, Th.; Andre, E.; Baldes, St.; Gebhard, P.; Klesen, M.; Kipp, M.; Rist, P.; Schmitt, M. (2004): A review on the development of embodied conversational agents and their application fields. In: Prendinger, H.; Ishizuka, M. (Eds): Life-Like Characters: Tools, Affective Functions, and Applications. Berlin et al.: Springer-Verlag, pp. 377-404.

Shneiderman, B. (1998): Designing the user interface: strategies for effective human-computer-interaction. Reading, MA et al.: Addison-Wesley.

Contact Information

PD Dr. Mandl / Matthias Görtz, University of Hildesheim
Marienburger Platz 22, D-31141 Hildesheim, Germany
Tel.: +49 (0) 5121/883 – 837
E-Mail: mandl@uni-hildesheim.de / matthiasgoertz@gmx.de