

# User Characteristics and Usage of Gesture and Speech in a Smart Office Environment

S. Schaffer, J. Seebode, I. Wechsung, F. Metze, and C. Kühnel

Quality and Usability Lab, TU Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany  
christine.kuehnel@telekom.de  
<http://www.qu.tlabs.tu-berlin.de>

**Key words:** usability studies, multimodal dialogue, haptic interfaces

## 1 Introduction

While interactions with applications via stylus or finger are wide-spread nowadays, their combination with speech or other input modalities might still improve the interfaces usability [1]. At Deutsche Telekom Laboratories (T-Labs), *attentive displays* present a wall-mounted interface for access to information about meetings and workplaces, room and workplace bookings, and general information. We enhanced the *attentive displays* (which were originally controllable by gesture only) with a speech recognition module, rendering the interface multimodal. In this paper, we present a first evaluation of this office information system, focusing on user characteristics. According to [2] gender differences in regard to attitude are expected as well as differences between expert and novice users (cp. [3]).

## 2 Method

In this study three different implementations of the system were compared, with the input modalities GESTURE, SPEECH and MULTIMODAL (gesture and speech) being varied as within factor. 17 male and 19 female German speakers ( $M_{age}=27.19$ ,  $SD_{age}=3.45$ ) took part of an one hour experiment. Half of the participants consisted of students. The other half – being T-Labs employees – were considered as **system experts** concerning the usage of the attentive displays. Demographic data was assessed with a pre-test questionnaire. The experience with speech-dialogue-systems (SDS) and with gesture-controlled-systems (GCS) were assessed as bipolar variables: 10 users had experience with SDS and 18 had experience with GCS, both groups are considered as **domain experts**.

The tasks to be accomplished by the participants consisted of changing screens, searching for the workplace of employees, for rooms and for bookings. The experiment was divided into three blocks. Each block consisted of 6 tasks that had to be carried out with the complying modality (GESTURE and SPEECH whose order was randomised, followed by MULTIMODAL input). Every block was followed by two questionnaires: a modified SASSI [4], adapted to gesture/multimodal input and the AttrakDiff [5].

### 3 Results and Analysis

*Intern vs. extern:* Neither the SASSI nor the AttrakDiff showed any differences between the user groups.

*Experienced vs. unexperienced users:* The AttrakDiff showed no differences. The SASSI exhibited differences for both domains: SDS experienced users rated the GESTURE input better on the scales accuracy, likeability and hability than those without experience. Regarding experience with gesture input experienced users rated the SPEECH input to be less accurate and the GESTURE input to be slower than those without experience.

*Gender differences:* On the AttrakDiff scale "hedonic quality - identity" female participants rated all three modalities better than men did. The SASSI yielded no differences between male and female participants.

### 4 Discussion

Our main findings are that (1) experience in operating the evaluated system had no influence on the user ratings, (2) ratings are effected by domain expertise in speech-dialog-systems and gesture-controlled-systems, (3) women can identify more strongly with the evaluated system than men – regardless of the input modality. Regarding (1) a possible explanation could be that system and domain expertise are confounded. Taking finding (2) into account it appears that system expertise is dominated by domain expertise. Gender differences were observed only for "hedonic quality - identity", and there seems to be no plausible explanation for this finding at the moment, so that more research is necessary. The analysis of domain expertise relies on answers to two questions only, conclusions drawn from this should be interpreted with care.

### References

1. Jokinen, K., Hurtig, T. 'User Expectations and Real Experience on a Multimodal Interactive System.' In Proceedings of the Interspeech 2006, Pittsburgh, US
2. Canada, K. and Brusca, F. 'The technological gender gap: Evidence and recommendations for educators and computer-based instruction designers.' In Educational Technology Research and Development 1991, 39 (2), pp. 43 –51
3. Kamm, C., Litman, D., Walker, M. 'From novice to expert: the effect of tutorials on user expertise with spoken dialogue systems.' In Proceedings of the Interspeech 1998, Sydney, Australia
4. Hone, K.S. and Graham, R. 'Towards a tool for the subjective assessment of speech system interfaces (SASSI).' In Natural Language Engineering 2000, 6(3/4), pp. 287–305
5. Hassenzahl, M., Burmester, M., Koller, F. 'AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität.' In Mensch & Computer 2003. Interaktion in Bewegung, pp. 187–196