Multimodality, Inclusive Design, and Intuitive Use

Anja B. Naumann
Deutsche Telekom Laboratories,
TU Berlin
Ernst-Reuter-Platz 7
D-10587, Berlin
anja.naumann@telekom.de

Ina Wechsung
Deutsche Telekom Laboratories,
TU Berlin
Ernst-Reuter-Platz 7
D-10587, Berlin
ina.wechsung@telekom.de

Jörn Hurtienne
Cambridge University,
Engineering Department,
Trumpington Street,
Cambridge CB2 1PZ, UK
j.hurtienne@eng.cam.ac.uk

ABSTRACT
With users being able to choose a modality matching their preferences and needs, multimodality is in line with the "Design for all"- principle. However, despite the potential advantages possible disadvantages are also reported: Multimodality requires more effort for combination and coordination of the modalities, and higher cognitive load may be the result. Thus, it is not clear if multimodality actually increases accessibility to information technology and if there are modalities that are more "inclusive" than others. Therefore the current study investigates if different user groups (older vs. younger users) benefit from new input modalities, such as touch, speed, and motion interaction, as well as their multimodal combination. It was shown that especially motion control does not correspond to older users' needs. Moreover, older users did not use the possibility of a flexible, multimodal interaction. The paper concludes with a discussion of the determinants of intuitive use in the light of the modalities investigated.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces And Presentation**]: User Interfaces: *interaction styles, user-centered design, voice I/O, haptic I/O, evaluation/methodology*

General Terms
Human Factors

Keywords
Multimodality, Age effects, Acessibility, Evaluation, Intuitive Use, Inclusive Design.

1. INTRODUCTION
Multimodal interfaces, similar to natural human-human communication, combine multiple sensory input and output channels [2]. This similarity leads to the expectation that multimodality in human-computer interaction provides more natural, robust and flexible interaction [3]. Moreover, with the users being able to choose a modality matching their preferences and needs, multimodality is in line with the design for all principle (e.g. [1][4]). However, despite these potential advantages also disadvantages are possible [1]: Multimodality requires more effort for combination and coordination of the modalities. Furthermore higher cognitive load may result [6], and intuitive use is less likely. Current research provides findings supporting both assumptions. Advantages [7] as well as disadvantages [8] are reported.

Thus, it is not clear if multimodality increases accessibility to information technology and if some modalities are more intuitive to use and also have a higher "inclusiveness" than others. The current study investigates if different user groups (older users vs. younger users) do equally benefit from new input modalities and multimodality.

2. METHOD
2.1 Participants
30 participants (15 male, 15 female) took part in the study. Half of them were younger than 35 years (M=29) and half of them were older than 55 years (M= 66). No one of the participants was familiar with either the application or the technical device.

2.2 Material
The device used was a smartphone (HTC Touch Diamond) controllable via motion (tilt and twist), speech and touch input (Figure 1). System output was graphical for all modalities. For motion control, additional tactile feedback (vibration) and for speech control, additional auditive feedback was given. The application tested was a multimodal mailbox system capable of handling speech-, e-mail- and fax-messages as well as forwarding calls and notifying of mailbox messages.

For measuring the consequences of intuitive use we used a questionnaire based on a definition of intuitive use as 'the subconscious application of prior knowledge that leads to effective interaction' [9]. This definition of intuitive use is captured in the subscales: *Perceived Cognitive Load* (internal consistency of the scale: Cronbach's alpha=0.90; 6 items, e.g. "The use of the system was not complicated."), *Perceived Achievement of Goals*: (Cronbach's alpha=0.95; 6 items, e.g. "I was able to reach all my goals with the system."), *Perceived Error Rate*: (Cronbach's alpha=0.94; 3 items, e.g. "No problems occurred while using the system."), *Perceived Effort of Learning*: (Cronbach's alpha=0.93; 6 items, e.g. "I immediately knew how to operate the system."), *Familiarity*: (Cronbach's alpha=0.93; 6 items, e.g. "The operation of the system always met my expectations."), *Global*: (1 item, "The use of the system was intuitive."). Items were answered on a five-point Likert scale (0= strongly disagree, 4= strongly agree). All results show scale means. Higher values indicate a higher probability for intuitive use of the system tested.

To assess the subjectively experienced effort, the SEA-scale [10] was employed. The scale ranges between 0 and 220 points, with higher values indicating higher perceived effort.

Furthermore, interaction data (successful task completion, aborts of task execution, task duration) were logged. In the multimodal test block the chosen modality was recorded.

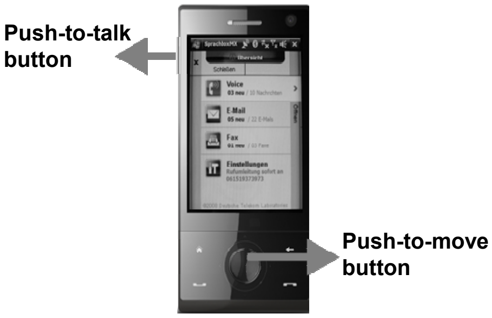

Figure 1. Tested prototype and allocation of relevant operating elements.

2.3 Procedure

There were four modality conditions. In each condition participants executed a total of 14 tasks (e.g. to get messages, reply to them, forward and sort messages, as well as changing notification options). If the goal was not achieved within three trials, task execution was aborted and the next task started. Afterwards, participants evaluated the interaction via the questionnaire. This was repeated for all three modalities – touch, speech, and motion. The sequence of the modalities was balanced between the participants. In the final condition, participants were free to choose the modalities they used for solving the task. Here, it was always possible to switch or to combine modalities. Again, the participants evaluated the interaction after solving all tasks in this condition.

3. RESULTS

3.1 Questionnaire Data

3.1.1 Intuitive Use

Regarding the questionnaire measuring consequences of intuitive use, differences between older and younger users were observed for the test condition with motion control and for the multimodal condition.

The condition motion control was rated worse by older users than by younger users on the scales *Perceived Cognitive Load,* t(28)=2.43, p=.022, d=.89, *Perceived Achievement of Goals,* t(28)= 4.27, p=.000, d=1.56 and *Perceived Effort of Learning,* t(28)=2.61, p=.014, d=.95. Also regarding the scale *Familiarity* older users tended to rate motion control poorer than younger users, t(28)=1.82, p=.080, d=.66.

In the multimodal condition differences occurred on the scale *Perceived Achievement of Goals,* again with worse ratings given by older users (t(27)= 2.51, p=.018, d= .92) .

3.1.2 Perceived Effort

Only in the touch condition, differences between older and younger users occurred (Figure 2): Younger users tended to rate the touch modality as more demanding than older users, t(15.32)=2.12, p=0.051, d=.77.

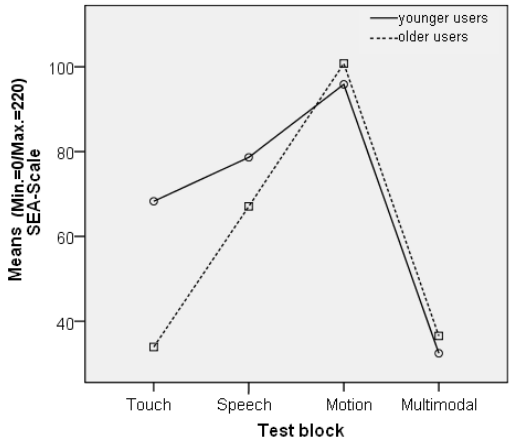

Figure 2. Ratings on SEA-scale for older and younger users by modality condition.

3.2 Interaction Data

3.2.1 Task Success and Task Aborts

Regarding successful first attempts, differences only occured in the condition with motion control (Figure 3), with the older users being less successful on the first attempt. Younger users solved more tasks in their first trial than older users did, t(28)=5.25, p=.000, d=1.92.

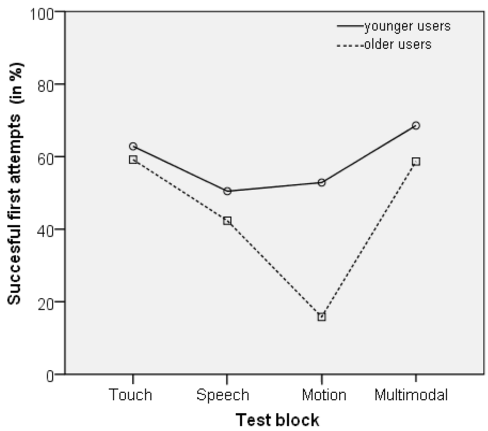

Figure 3. Percentages of successful first attempts for both age groups by modality condition.

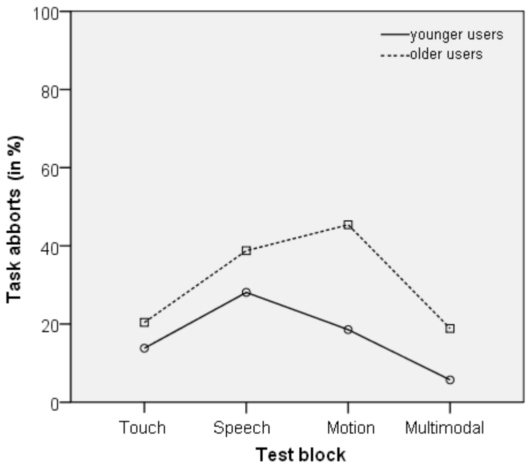

Figure 4. Percentages of aborts for both age groups by modality condition.

Task abortions occurred more frequently with older users in all conditions, except touch: speech, $t(28)=2.61$, $p=.014$, $d=.95$, motion, $t(28)=3.34$, $p=.002$, $d=1.21$, and multimodal, $t(27)=3.73$, $p= 0.001$, $d= 1.36$ (Figure 4).

3.2.2 Task Duration

In all conditions, longer task durations were observed for older users (s. Table 1).

Table 1. Task duration in minutes for both age groups by modality condition

Modality	Older users		Younger users		t(df)	p	d
	M	SD	M	SD			
Touch	06:25	01:36	05:01	01:12	2.70 (28)	.012	.99
Motion	12:07	04:20	07:46	02:10	3.47 (28)	.002	1.27
Speech	07:11	00:51	05:57	01:11	3.23 (28)	.003	1.18
Multimodal	05:21	01:05	03:44	00:55	4.39 (27)	.000	1.63

3.2.3 Modality Usage and Modality Switches

Regarding the preference for any modality in the multimodal condition, no differences between younger and older users were observed. Over all tasks, touch was the preferred modality in the multimodal test block for both groups.

Regarding modality switches, older users showed a less flexible interaction strategy than younger users (Figure 5):

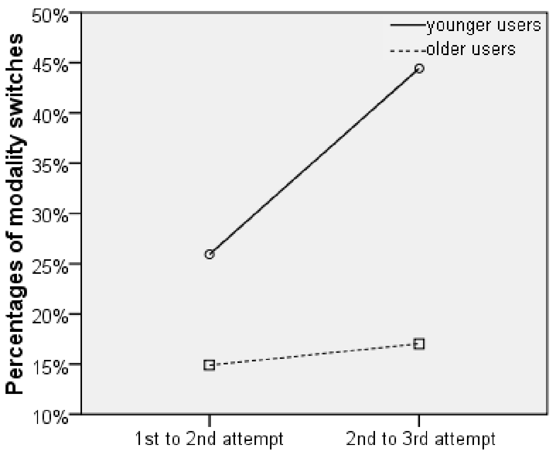

Figure 5. Percentages of modality switches for both age groups.

After a failed attempt, older users were less likely to switch the modality than younger users (second attempt: $\chi^2(1, N=147)=4.27$, $p=.044$, third attempt: $\chi^2(1, N=74)=6.54$, $p=.015$).

4. DISCUSSION AND CONCLUSION

The results show differences in subjective ratings of intuitive use mainly for motion control. Older users rated this modality worse than younger users did. This is line with the interaction data: Older users were less successful when using motion control than younger users. Thus motion seems to be the modality least appropriate for older users.

In line with previous research [11][12] the interaction data showed a generally lower performance for the older users through all test blocks. Furthermore older participants used the flexibility offered with the multiple input modalities to a lesser extent than younger users did. However, regarding the subjective data only few differences between the age groups were observed in the modalities different from motion control.

In summary it was shown, that for the tested system the motion control was the only modality not suitable for older users. An explanation might be the age related motor impairments known from previous research (e.g.[13]). For all other modalities, subjective ratings were predominantly as good as, or in the touch condition even better than the ratings from the younger users. Nevertheless, the performance data was worse for older users and one of the often mentioned advantages of multimodality, the possibility of more flexible interaction, seems not to apply to older users.

What consequence for the ongoing discussion of intuitive use in inclusive design can be drawn? If we follow the definition of intuitive use as ‘the subconscious application of prior knowledge that leads to effective interaction’, then we must ask how the different aspects of the definition have been met by the data in our study. The requirement of effective interaction is met by the performance data (successes and abortions), as well as the subscales *Perceived Error Rate* and *Perceived Achievement Of Goals* of the questionnaire. Another entailment of the definition is that intuitive interaction should result in lower mental effort. This was shown in the data on subjective effort and in the subscales *Perceived Cognitive Load* and *Perceived Effort Of Learning* of the intuitive use questionnaire.

The most interesting consequence refers to the concept of prior knowledge. It could be argued that the participants (old and young) relied on their prior knowledge of direct manipulation that was exploited most in the condition with touch control. There also is no great difference in the required knowledge, whether physical or virtual buttons are touched. The knowledge gap to the other modalities seems greater, as most of the participants were not familiar with speech or motion control. Therefore, the touch modality achieved the highest ratings and yielded the best performance, compared to other modalities.

But is it really just the familiarity with the style of interaction? With speech and motion control the interaction was rather awkward, because a button had to be pushed while speaking to the phone or applying motion control. An argument could therefore be made, that it was not poor prior knowledge that yielded lower intuitive use of these modalities, but that it was poorly designed motor interaction. Intuitive interaction could therefore to a far larger extent depend on hardware ergonomics issues than suggested. Our data, however, cannot resolve the conflict between prior experience and ergonomic design as the determinants of intuitive use. Further research needs to be conducted to disentangle the cognitive from the sensorimotor aspects of interaction. Only then can a definition of intuitive use be verified that solely relies on prior knowledge.

5. REFERENCES

[1] Jokinen, K. & Raike A. 2003. Multimodality - technology, visions and demands for the future. The 1st Nordic Symposium on Multimodal Interfaces, Copenhagen, September, 2003.

[2] Chen, F. 2006. Designing Human Interface in Speech Technology. Springer, Berlin.

[3] Oviatt, S. 1999. Ten myths of multimodal interaction. Communications of the ACM, 42 (11), 74-81.

[4] Oviatt, S. L. 2003. Multimodal Interfaces. In J. A. Jacko & A. Sears (Eds.), The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications. Lawrence Erlbaum Associates, Mahwah, 286–304.

[5] Oviatt, S., & Wahlster, W. 1997. Introduction to This Special Issue on Multimodal Interfaces Multimodal Interfaces. Human-Computer Interaction, 1&2, 1–5.

[6] Schomaker, L., Nijtmans, J., Camurri, A., Lavagetto, F.,Morasso, P., Benoît, C., Guiard-Marigny, T., Le Goff, B., Robert-Ribes, J., Adjoudani, A., Defée, I., Münch, S., Hartung, K. & Blauert, J. 1995. A Taxonomy of Multimodal Interaction in the Human Information Processing System. A Report of the ESPRIT Project 8579 MIAMI. NICI, Nijmegen.

[7] Emery, V. K., Edwards, P. J., Jacko, J. A., Moloney, K. P., Barnard, L., Kongnakorn, T., Sainfort, F., and Scott, I. U. 2003. Toward achieving universal usability for older adults through multimodal feedback. In Proceedings of the 2003 ACM Conference on Universal Usability, 46-53.

[8] Wechsung, I. & Naumann, A. B. 2008. Evaluation Methods for Multimodal Systems. A Comparison of Standardized Usability Questionnaires. In E. André, L. Dybkjær, W. Minker, H. Neumann, R. Pieraccini, and M. Weber (Eds.), Proceedings of PIT 4th IEEE Tutorial and Research Workshop on Perception and Interactive Technologies For Speech-Based Systems: Perception in Multimodal Dialogue Systems PIT. LNAI, Vol. 5078. Springer, Berlin, 276-284.

[9] Mohs, C., Hurtienne, J., Scholz, D., & Rötting, M. 2006. Intuitivität: definierbar, beeinflussbar, überprüfbar! [Intuitiveness: definable, influenceable, verifiable!] In VDI / VDE-Gesellschaft Mess- und Automatisierungstechnik (Eds.) Useware 2006. VDI-Verlag, Düsseldorf, 215-224.

[10] Eilers, K., Nachreiner, F., & Hänecke, K. 1986. Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung [Development and validation of a scales measuring subjective perceived effort]. Zeitschrift für Arbeitswissenschaft, 40, 215-224.

[11] Echt, K. V., Morrell, R. W., & Park, D. C. 1998. Effects of age and training formats on basic computer skill acquisition in older adults. Educational Gerontology, 24(1), 3-25.

[12] Lewis, T., Langdon, P. M., & Clarkson, P. J. 2008. Prior experience of domestic microwave cooker interfaces: A user study. In P. M. Langdon, P. J. Clarkson, & P. Robinson (Eds.), Designing Inclusive Futures. Springer, London. 95–106.

[13] Walker, N., Philbin, D. & Fisk, D. 1997. Age related differences in movement control: Adjusting sub movement structure to optimize performance. Journal of Gerontology: Psychological Science, 52B (1), 40-52.