

Multimodale Anwendungen: Was beeinflusst die Wahl der Modalität?

Anja Naumann, Ina Wechsung und Sebastian Möller

Deutsche Telekom Laboratories, Technische Universität Berlin

1 Einleitung

Da die zwischenmenschliche Kommunikation mehr als eine Modalität beinhaltet, wird angenommen, dass multimodale Schnittstellen natürlicher seien als unimodale (Kallinen & Ravaja 2005). Weiterhin wird erwartet, dass multimodale Schnittstellen flexibler, effizienter und robuster sind als unimodale (Oviatt 1999). Andererseits kann Multimodalität auch die kognitive Belastung erhöhen, da die Entscheidung für die angemessene Modalität zusätzliche kognitive Ressourcen verbraucht, und beim Einsatz mehrerer Modalitäten interferieren diese möglicherweise auch miteinander (Schomaker et al. 1995).

2 Der Einfluss von Aufgabe und Kontext auf die Wahl der Modalität

Obwohl Untersuchungen zeigen, dass Benutzer multimodale Interaktion beispielsweise für räumliche Aufgaben bevorzugen (Oviatt 1999) und deutliche Präferenzen für bestimmte Modalitäten in Abhängigkeit von der Aufgabe aufweisen (Ren et al. 2000), bleibt unklar, ob man diese Ergebnisse auch auf andere Aufgaben übertragen kann. Der potentielle Mehrwert von Multimodalität hängt demnach zumindest von folgenden Faktoren ab: der Aufgabe, der Situation und den zur Auswahl stehenden Modalitäten.

Ziel einer empirischen Studie (Wechsung, Naumann & Möller 2008) war zu untersuchen, ob Benutzer von angebotener Multimodalität Gebrauch machen und unter welchen Umständen sie dies tun. Des Weiteren wurde überprüft, ob die Präferenz für eine Modalität, die ein Benutzer angibt, auch mit der tatsächlichen Nutzung dieser Modalität übereinstimmt.

Der *MediaScout*, eine multimediale Web-basierte Anwendung zur Auswahl und zum Managen von Medieninhalten, wurde auf zwei multimodalen Geräten installiert: 1. PDA (Fujitsu-

Siemens Pocket LOOX T830), 2. Ultramobile Tablet PC (Samsung Q1-Pro 900 Casomii), 3. Desktop-PC (Kontrollbedingung). Die ersten beiden Geräte sind multimodal und sowohl über eine graphische, berührungssensitive Benutzungsoberfläche (Touch-Screen) als auch per Sprache steuerbar. Beim PDA war zusätzlich Bewegungssteuerung möglich, z.B. konnte eine Menüauswahl durch Drehen/Kippen des Gerätes getroffen werden. Die Bedienung des unimodalen PCs erfolgte mittels Maus und Tastatur.

21 Probanden hatten eine Serie von fünf verschiedenen Aufgabentypen mit jedem Gerät auszuführen, z.B. eine Option aus einem Drop-Down-Menü auswählen oder eine Telefonnummer eingegeben. Diese Aufgaben sollten mit einer vorgegebenen Modalität ausgeführt werden. Dies wurde für jede von dem jeweiligen Gerät unterstützte Modalität wiederholt. Danach wurden die Aufgaben wieder präsentiert, die Probanden konnten nun jedoch für ihre Bearbeitung eine Modalität frei wählen. Danach füllten die Probanden verschiedene Fragebögen zur Beurteilung des gerade getesteten Gerätes, z.B. SUMI (Kirakowski & Corbett 1993), aus. Dieses Vorgehen wurde für jedes der drei Geräte wiederholt. Nach dem letzten Gerät füllten die Probanden einen abschließenden Fragebogen zum Gesamteindruck und zu den Modalitätenpräferenzen aus. Zusätzlich wurden Log-Dateien sowie Video- und Audio-dateien während der gesamten Untersuchung aufgezeichnet.

Modalitätennutzung. Bei beiden Geräten wurden die meisten Aufgaben mittels Touchscreen (GUI) ausgeführt (siehe Abbildung 1). Eine detailliertere Analyse zeigt, dass die Eingabe einer Telefonnummer die einzige Aufgabe war, bei der die Spracheingabe etwa genauso häufig verwendet wurde wie die GUI-Eingabe. Bei allen anderen Aufgaben war das GUI deutlich die erste Wahl. Unerfahrene Benutzer von Informations- und Kommunikationstechnologie verwendeten die Bewegungssteuerung des PDA häufiger als erfahrene Benutzer und Frauen häufiger als Männer.

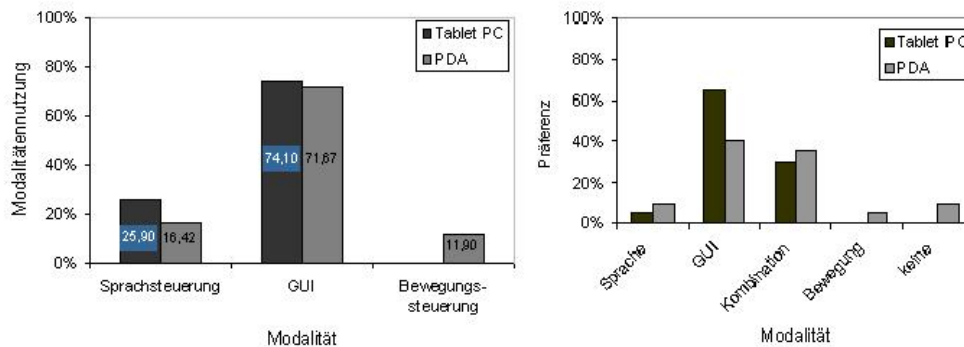


Abbildung 1: links- Modalitätennutzung über alle Aufgaben bei der Interaktion mit Tablet PC und PDA, rechts- bevorzugte Modalitäten für Tablet PC und PDA (Bewegungssteuerung nur im PDA implementiert)

Modalitätenpräferenz. Im abschließenden Fragebogen wurden die von den Probanden präferierten Modalitäten erfragt. Dabei konnten sowohl einzelne Modalitäten als auch Modalitätenkombinationen gewählt werden. Konsistent mit den Log-Daten bevorzugten die meisten Probanden das GUI vor der Kombination von GUI und Sprachsteuerung und vor Sprachsteuerung allein (siehe Abbildung 1). Beim PDA war die Kombination von Bewegungssteuerung und GUI die Modalitätenkombination, die von den meisten Probanden bevorzugt wurde.

Subjektive Beurteilung. Der PDA, das Gerät mit den meisten Modalitäten, wurde am besten beurteilt, der Tablet PC am zweitbesten und der unimodale PC am schlechtesten. Unterschiede zwischen erfahrenen und unerfahrenen Benutzern bzw. Frauen und Männern ergaben sich nicht.

Zusammenfassend ist festzustellen, dass in erster Linie die Charakteristika der Aufgabe zu bestimmen scheinen, ob eine angebotene Modalität genutzt wird oder nicht. Die Benutzer wählen die Modalität, die am effizientesten zur Zielerreichung führt. Beispielsweise konnte die Telefonnummer-Eingabe mittels Sprachsteuerung mit weniger Interaktionsschritten ausgeführt werden als mittels GUI. Ausschließlich bei dieser Aufgabe wurde die Sprache ebenso häufig verwendet wie die GUI-Eingabe. Das GUI wurde hier zwar trotz geringerer Effizienz noch vergleichbar oft genutzt, dies kann jedoch mit der hohen Vertrautheit der Probanden mit dieser Eingabemöglichkeit begründet werden. Möglicherweise erhöht sich der Gebrauch von Sprachsteuerung mit zunehmender Übung. Die Nutzung der Modalitäten war konsistent mit der angegebenen Modalitätenpräferenz: Die Mehrzahl der Aufgaben wurde mit dem GUI ausgeführt, das ebenfalls als die präferierte Modalität angegeben wurde. Interessanterweise beeinflussen die angebotenen Modalitäten die subjektive Beurteilung selbst wenn sie selten genutzt werden: Der PDA, das Gerät mit der größten Auswahl an Modalitäten, wurde am besten bewertet, obwohl für fast alle Aufgaben das GUI die bevorzugte Eingabemöglichkeit war.

3 Verschiedene Benutzergruppen und Multimodalität

Des Weiteren ist zu klären, ob verschiedene Benutzergruppen (z.B. ältere vs. jüngere Benutzer) gleich gut von neuen Modalitäten und Multimodalität profitieren. Dafür spricht die Möglichkeit des Benutzers, eine Modalität zu wählen, die seinen Präferenzen und Bedürfnissen entspricht, was mit dem Prinzip des barrierefreien Designs in Einklang (Oviatt 2003, Jokinen & Raike 2003) ist. Dagegen spricht die mögliche zusätzliche kognitive Zusatzbelastung (Schomaker et al. 1995). Die Ergebnisse empirischer Studien zeigen bisher sowohl Vorteile (Emery et al. 2003) als auch Nachteile (Wechsung & Naumann 2008) auf.

In einer weiteren Studie (Hurtienne, Wechsung & Naumann 2009) haben wir daher untersucht, ob Multimodalität den Zugang zu Informationstechnologie erhöht und ob manche Modalitäten intuitiver zu benutzen sind als andere. Das verwendete Gerät war ein Smartphone (HTC Touch Diamond), steuerbar mittels Bewegungssteuerung (Drehen/Kippen), Sprache

und Eingabe über eine berührungssensitive Oberfläche (Touch). Die Systemausgabe war für alle Modalitäten graphisch. Für die Bewegungssteuerung wurde zusätzlich taktile Rückmeldung gegeben und für die Sprachsteuerung zusätzlich auditive Rückmeldung. Die getestete Anwendung war ein multimodales Mailbox-System, mit dem Sprach-, E-mail- und Fax-Nachrichten verwaltet sowie Anrufe weitergeleitet und Benachrichtigungen bei einkommenden Nachrichten verschickt werden können.



Abbildung 2: Getesteter Prototyp und Lage der relevanten Steuerelemente

30 Probanden hatten in vier unterschiedlichen Modalitätsbedingungen (Touch, Sprache, Bewegung und freie Modalitätswahl) jeweils 14 Aufgaben zu bearbeiten (z.B. Abrufen von Nachrichten, Antworten, Weiterleiten und Sortieren, Ändern von Benachrichtigungsoptionen). Im Anschluss nach jeder Bedingung beurteilten die Probanden die Interaktion mittels eines Fragebogens zu intuitiver Interaktion.

Die Ergebnisse zeigen Unterschiede in der subjektiven Beurteilung der intuitiven Benutzung hauptsächlich für die Bewegungssteuerung. Ältere Benutzer (älter als 55 Jahre) beurteilten diese Modalität schlechter als jüngere Benutzer (jünger als 35 Jahre). Dies deckt sich auch mit den Interaktionsdaten: Ältere Benutzer waren weniger erfolgreich als jüngere Benutzer. Demnach scheint Bewegungssteuerung die am wenigsten geeignete Modalität für ältere Benutzer zu sein. Eine Erklärung hierfür könnten die altersbedingten motorischen Einschränkungen sein, die in der Literatur berichtet werden (z.B. Walker et al. 1997). Für alle anderen Modalitäten gab es im Wesentlichen keine Unterschiede zwischen jüngeren und älteren Benutzern hinsichtlich der subjektiven Beurteilungen.

Literaturverzeichnis

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*, S. 46-53.

- Hassenzahl, M., Burmester, M. & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: J. Ziegler; G. Szwillus (Hrsg.): *Mensch & Computer 2003. Interaktion in Bewegung*. Stuttgart: Teubner, S. 187-196.
- Hurtienne, J., Wechsung, I. & Naumann, A.B. (2009). *Multimodality, Inclusive Design, and Intuitive Use*. (submitted manuscript).
- Jokinen, K. & Raike A. (2003). Multimodality - technology, visions and demands for the future. *The 1st Nordic Symposium on Multimodal Interfaces*, Copenhagen.
- Kallinen, K. & Ravaja, N. (2005). Effects of the Rate of Computer-Mediated Speech on Emotion-related Subjective and Physiological Responses. *Behaviour & Inf. Technology*, 24, S. 365 – 373.
- Kirakowski, J. & Corbett, M. (1993). SUMI: The Software Usability Measurement Inventory. *British Journal of Education Technology*, 24, 3, S. 210-212.
- Oviatt, S. L. (1999). Ten myths of multimodal interaction. *Communications of the ACM*, 42, 11. S. 576-583.
- Oviatt, S. L. (2003). Multimodal Interfaces. In J. A. Jacko & A. Sears (Hrsg.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. Mahwah: Lawrence Erlbaum Associates, S. 286–304.
- Ren, X., Zhang, G. & Dai, G. (2000). An Experimental Study of Input Modes for Multimodal Human-Computer Interaction. In: T. Tan, Y. Shi, & W. Gao (Hrsg.). *Lecture Notes in Computer Science, Vol. 1948*. London: Springer, S. 49-56.
- 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. A. (1995). *Taxonomy of Multimodal Interaction in the Human Information Processing System*. A Report of the ESPRIT Project 8579 MIAMI. NICI, Nijmegen.
- 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.
- Wechsung, I. & Naumann, A. (2008). Evaluation Methods for Multimodal Systems: A Comparison of Standardized Usability Questionnaires. In E. André, L. Dybkjær, W. Minker, H. Neumann, R. Pieraccini & M. Weber (Hrsg.), *Proceedings of 4th IEEE Workshop on Perception and Interactive Technologies for Speech-based Systems (PIT08)*, S. 276-284. Heidelberg: Springer.
- Wechsung, I., Naumann, A., & Möller, S. (2008). Multimodale Anwendungen: Einflüsse auf die Wahl der Modalität. *Mensch & Computer 2008*. S. 437-440. München: Oldenbourg Wissenschaftsverlag.

Kontakt

anja.naumann@telekom.de