

# Multimodale Anwendungen: Einflüsse auf die Wahl der Modalität

Ina Wechsung, Anja Naumann, Sebastian Möller

Deutsche Telekom Laboratories, Technische Universität Berlin

## Zusammenfassung

In dieser Evaluationsstudie wurde ein Web-basiertes Medienmanagementsystem auf zwei multimodalen und einem unimodalen Gerät zur Lösung von Aufgaben eingesetzt. Ziel war es herauszufinden, ob Benutzer Gebrauch von angebotener Multimodalität machen und unter welchen Umständen sie dies tun. Weiterhin wurde untersucht, ob die von den Benutzern angegebene Präferenz für eine Modalität auch mit ihrer tatsächlichen Nutzung übereinstimmt. Die Ergebnisse zeigen, dass die Nutzung einer Modalität sowohl von der auszuführenden Aufgabe als auch von der Effizienz der jeweiligen Modalität im Hinblick auf die Erreichung des Aufgabenziels abhängig ist. Sowohl die eigentliche Nutzung als auch das Angebot an Modalitäten scheinen die subjektiven Benutzerurteile zu beeinflussen.

## 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). 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 der vorliegenden Studie ist daher 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.

## 2 Methode

*Probanden.* An der Studie nahmen 11 Männer und 10 Frauen teil (mittleres Alter: 31,2 Jahre). Für die Teilnahme erhielten sie einen Büchergutschein. Etwa die Hälfte der Probanden waren erfahrene Benutzer von Informations- und Kommunikationstechnologie (IKT), die andere Hälfte unerfahrene IKT-Benutzer, die IKT-Erfahrung wurde mittels eines Screening-Fragebogens bei der Rekrutierung erhoben.

*Anwendung und Geräte.* 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 Benutzungsoberfläche mit einem 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.

*Aufgaben und Ablauf.* Die Probanden hatten eine Serie von fünf verschiedenen Aufgabentypen mit jedem Gerät auszuführen: Navigation (7 Aufgaben); Felder ankreuzen oder Kreuze entfernen (6 Aufgaben); Option aus einem Drop-Down-Menü auswählen (4 Aufgaben); Taste drücken (3 Aufgaben) und Telefonnummer eingeben (1 Aufgabe). Der Test wurde in einem Usability-Labor durchgeführt. Zunächst erklärte der Versuchsleiter das erste Gerät und die Anwendung *MediaScout*, dann sollten die Probanden die o.g. Aufgaben mit einer vorgegebenen Modalität ausführen. 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 die Fragebögen zur Beurteilung des gerade getesteten Gerätes aus. Dieses Vorgehen wurde für jedes der drei Geräte wiederholt, die Abfolge der Geräte war randomisiert. Nach dem letzten Gerät füllten die Probanden einen abschließenden Fragebogen zum Gesamteindruck und zu den Modalitätenpräferenzen aus.

*Tests.* Verschiedene Fragebögen, darunter der SUMI (Kirakowski & Corbett 1993), wurden eingesetzt, um unterschiedliche Aspekte der Benutzereindrücke zu erfassen. Zusätzlich wurden Log-Dateien und Video- und Audiodaten während der gesamten Untersuchung aufgezeichnet. Hier werden im Wesentlichen die Ergebnisse des SUMI, des abschließenden Fragebogens zu den Modalitäten und zum Gesamteindruck und die Log-Dateien berichtet. Um zu analysieren, welche Modalität die Probanden benutzen wollten, wurden die Log-Daten für den Testblock, in dem die Modalität frei gewählt werden konnte, annotiert. Für jede Aufgabe wurde die Modalität, die zuerst zu ihrer Lösung benutzt wurde, für die weitere Analyse ausgewählt. Daraus sind die Prozentsätze für die Modalitätennutzung pro Aufgabe errechnet worden.

### 3 Ergebnisse und Diskussion

*Modalitätenutzung.* In Abbildung 1 sind die bei jedem multimodalen Gerät benutzten Modalitäten gemittelt über alle Aufgaben dargestellt. Bei beiden Geräten wurden die meisten Aufgaben mittels Touchscreen (GUI) ausgeführt. Die Unterschiede zwischen den einzelnen Modalitäten sind sowohl für den Tablet-PC (Wilcoxon:  $Z = -2,38$ ,  $p = .003$ ) als auch für den PDA signifikant (Friedman:  $\chi^2(2) = 18,88$ ,  $p = 0.000$ ). Eine detailliertere Analyse zeigt, dass die Eingabe einer Telefonnummer die einzige Aufgabe war, bei der die Spracheingabe (Tablet-PC 52%, PDA 33%) etwa genauso häufig verwendet wurde wie die GUI-Eingabe (Tablet-PC 48%, PDA 67%). Bei allen anderen Aufgaben war das GUI deutlich die erste Wahl. Unerfahrene IKT-Benutzer ( $M=20\%$ ,  $SD=16,8$ ) verwendeten die Bewegungssteuerung des PDA häufiger als erfahrene IKT-Benutzer ( $M=6\%$ ;  $SD=10,5$ ;  $U=22,00$ ,  $p=0.032$ ) und Frauen ( $M=17\%$ ,  $SD=12.2$ ) häufiger als Männer ( $M=7\%$ ,  $SD=16,0$ ;  $U= 27,50$ ,  $p=0.045$ ).

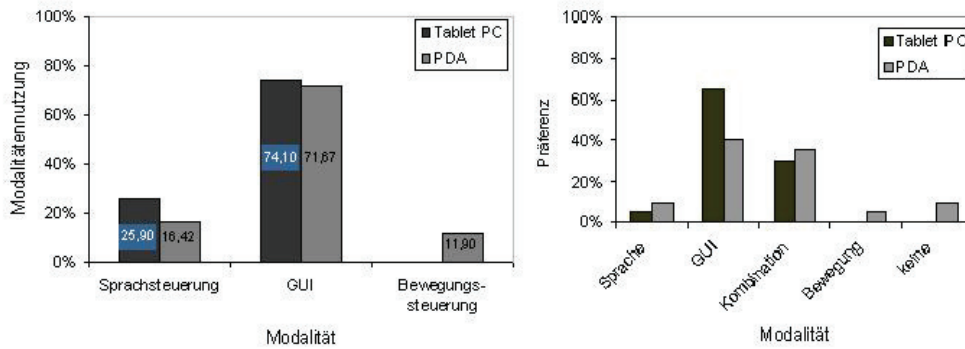


Abbildung 1: links- Modalitätenutzung über alle Aufgaben bei der Interaktion mit Tablet-PC und PDA, rechts- bevorzugte Modalitäten für Tablet-PC und PDA (Bewegungssteuerung nur im PDA implementiert)

*Modalitätenpräferenz.* Im abschließenden Fragebogen wurden die von den Probanden präferierten Modalitäten erfragt. Dabei konnten sowohl einzelne Modalitäten als auch Modalitätenkombinationen gewählt werden. Die Angabe, dass keine der angegebenen Modalitäten bevorzugt wird, war ebenfalls möglich. Konsistent mit den Log-Daten bevorzugten die meisten Probanden das GUI vor der Kombination von GUI und Sprachsteuerung und vor Sprachsteuerung allein,  $\chi^2(2) = 10,90$ ,  $p = 0.004$  (siehe Abbildung 1). Beim PDA war die Kombination von Bewegungssteuerung und GUI die Modalitätenkombination, die von den meisten Probanden bevorzugt wurde. Zwei Probanden berichteten keine Präferenz für eine Modalität.

*Subjektive Beurteilung.* Die Analyse der Rohdaten der globalen Skala des SUMI-Fragebogens zeigte signifikante Unterschiede,  $F(2,40)=6,56$ ,  $p=.003$ ; part.  $\eta^2=.247$ : Der PDA, das Gerät mit den meisten Modalitäten, wurde am besten beurteilt ( $M=45,29$ ,  $SD=10,14$ ), der Tablet-PC am zweitbesten ( $M=40,19$ ,  $SD=7,87$ ) und der unimodale PC am schlechtesten ( $M=38,04$ ,  $SD=7,06$ ). Unterschiede zwischen IKT-erfahrenen und IKT-unerfahrenen IKT-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. Diese Diskrepanz zwischen Nutzungsverhalten und Bewertung soll in einem nächsten Schritt analysiert werden. Dazu werden die gegenwärtigen Ergebnisse den Daten aus den zusätzlich eingesetzten Fragebögen (u. a. AttrakDiff, Hassenzahl et al. 2003) gegenübergestellt. Weiterhin ist geplant, den Einfluss von Aufgabenbearbeitungszeit und Fehlerrate auf die Modalitätenwahl zu untersuchen.

### Literaturverzeichnis

- 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.
- 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.
- 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 (Eds.). *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.

### Kontakt

ina.wechsung@telekom.de