

Qualitätsintegration bei multimodaler Interaktion

Sebastian Möller, Benjamin Weiss, Ina Wechsung, Christine Kühnel, Klaus-Peter Engelbrecht

Quality and Usability Lab, Deutsche Telekom Laboratories, TU Berlin, E-Mail: sebastian.moeller@telekom.de

Einleitung

Gestalter von multimodalen Applikationen gehen bislang davon aus, dass das Verfügbarmachen verschiedener Interaktionsmodalitäten die vom Nutzer erfahrene Qualität steigern wird. Dennoch gibt es abgesehen vom bekannten Zeige-Szenario „Put that there“ wenig überzeugende Beispiele vom erfolgreichen simultanen Einsatz verschiedener Modalitäten. Andererseits können verschiedene Modalitäten Teilaspekte der Qualität durchaus positiv beeinflussen, wie bspw. die Effektivität oder die Effizienz, und sie erhöhen die Nähe zu einer natürlichen menschlichen Kommunikation.

In diesem Beitrag werden theoretische Überlegungen darüber angestellt, auf welchen Ebenen eine Integration unterschiedlicher Modalitäten für den Nutzer gewinnbringend sein kann. Grundlage hierzu ist eine Taxonomie von Qualitätsaspekten multimodaler Interaktion. Zu den einzelnen Aspekten geben wir Messmethoden an und zitieren Beispiele der Beeinflussung unterschiedlicher Interaktionsmodalitäten. Wir schließen mit einer Reihe offener Fragen, deren Beantwortung die Gestaltung multimodaler Systeme erleichtern würde.

Taxonomie von Qualitätsaspekten

Die Qualität multimodaler Mensch-Maschine-Interaktion

kann von zwei Seiten beleuchtet werden. Systementwickler beschäftigen sich zumeist mit der Performanz ihrer Systeme, d.h. wie gut die Systeme die gewünschten Funktionen erfüllen. Metriken der Performanz lassen sich für den gesamten Interaktionszyklus zwischen Mensch und Maschine auf verschiedenen Ebenen definieren, wie es in der oberen Hälfte der Abb. 1 dargestellt ist.

Für den Nutzer zählt aber vielmehr die Qualität, d.h. das Ergebnis eines Wahrnehmungs- und Beurteilungsprozesses, bei dem der Nutzer das Erfahrene mit dem Erwarteten oder Erwünschten vergleicht. Qualität kann wiederum auf unterschiedlichen Ebenen definiert werden, wie der untere Teil von Abb. 1 zeigt.

Sowohl die Performanz als auch die Qualität kann vom Einsatz zusätzlicher Modalitäten bei der Interaktion profitieren. Im Folgenden werden wir einzelne Punkte der Taxonomie beleuchten, an denen ein Gewinn zu erwarten ist.

Performanz

Besonders augenfällig ist der mögliche Zugewinn bei der Leistung der Eingabekomponente (*Input Performance*): So können bspw. audiovisuelle Erkener insbesondere in geräuschbehafteten Umgebungen bessere Erkennungsraten erzielen als reine Spracherkener, und Gestenerkener auf einem berührungssensitiven Bildschirm können nochmals

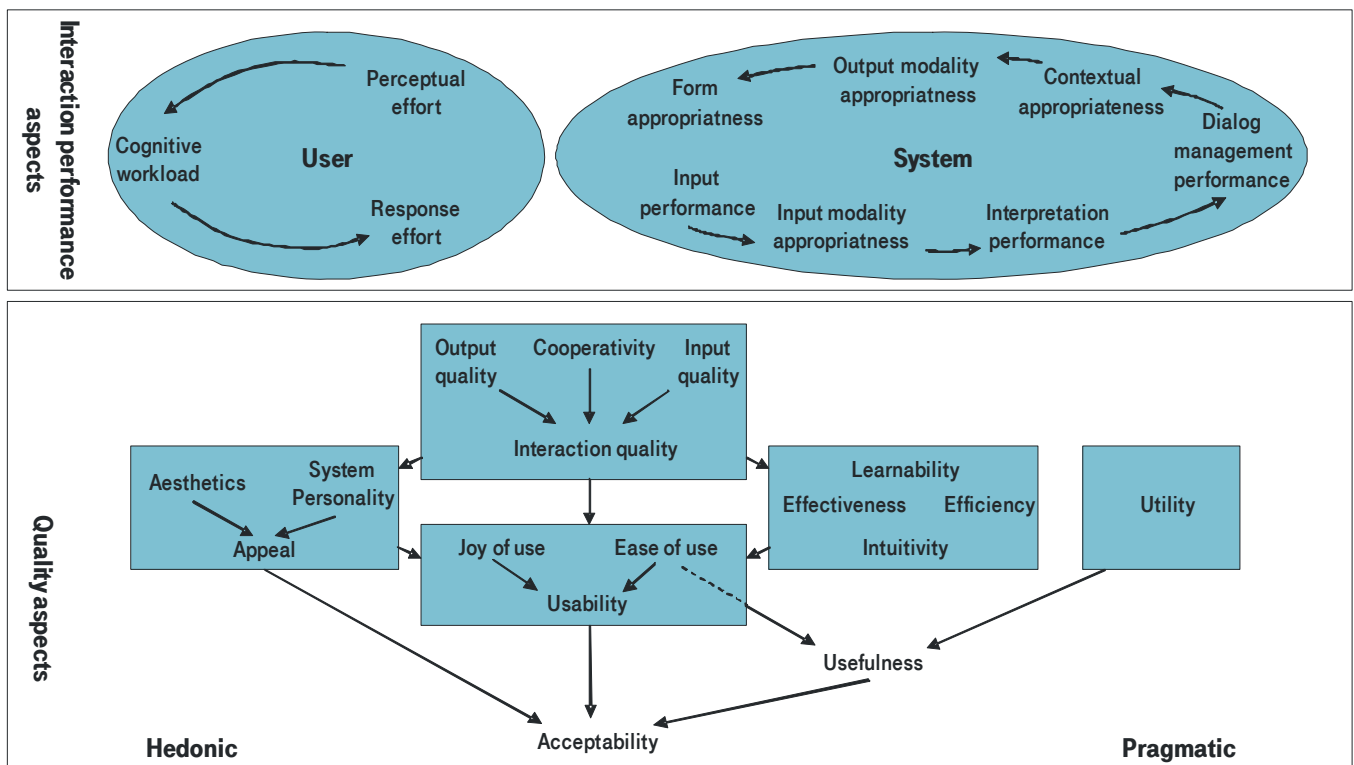


Abbildung 1: Taxonomie der Qualität multimodaler Interaktion, aus [1]. Oben: Aspekte der Interaction Performance; unten: Aspekte der Qualität.

höhere Erkennungsraten erreichen. Der Zugewinn an Erkennungsleistung hängt meist vom Benutzungsumfeld sowie teilweise auch von Nutzereigenschaften ab (*Modality Appropriateness*), und lässt sich normalerweise in Form von Fehlerraten der beteiligten Erkennermodule quantifizieren. Die Fusion der erkannten Informationen lässt sich anschließend auf semantischer Ebene als Konzeptfehlerraten quantifizieren (*Interpretation Performance*), welche sich ebenfalls durch eine geschickte Kombination von Eingabemodalitäten reduzieren lassen sollte. Dabei ist zu berücksichtigen, dass Modalitäten redundant oder komplementär sein können, und dass sie zeitlich sequenziell oder parallel verwendet werden können [2].

Auch auf der System-Ausgabeseite lassen sich durch geschickte Kombination von Modalitäten Vorteile erzielen. Dabei sollten die Eigenschaften von Modalitäten berücksichtigt werden, wie sie z.B. von Bernsen und Dybkjær [3] klassifiziert wurden, wie auch ihre inhaltliche und zeitliche Verwendung (s.o.). Auch innerhalb jeder Modalität muss die Form adäquat gewählt werden, sodass die Informationen bestmöglich präsentiert werden können. Dies lässt sich durch den Erfolg bei der Übermittlung der Informationen quantifizieren (bspw. Verständlichkeit/Verstehbarkeit).

Bezüglich des Effekts von Multimodalität auf die Nutzung kognitiver Ressourcen existieren verschiedene Theorien. Nach Wickens [4] kann durch die Belegung unterschiedlicher kognitiver Stadien (Perzeption, Kognition, Antwort) und verschiedener Antwortebenen (räumlich/manuell vs. verbal) durch unterschiedliche Sinne die kognitive Belastung reduziert werden. Allerdings wurden auch gegenteilige Effekte beobachtet.

Qualität

Die Verwendung unterschiedlicher Modalitäten beeinflusst natürlich direkt die Interaktionsqualität, bspw. wenn die Eingabequalität oder die Ausgabequalität des Systems vom Nutzer als unzureichend wahrgenommen wird. Der Nutzer wird dann bestimmte Verhaltensmuster entwickeln, welche Modalitäten er in welchen Situationen verwendet. Hinzu kommen persönliche Präferenzen, die natürlich auch durch die Erfahrung beeinflusst werden.

Einen leicht messbaren Nutzen erzielt Multimodalität bei der Verbesserung der Effektivität (Grad der Zielerreichung) und/oder der Effizienz (Einsatz von Ressourcen im Bezug zum Grad der Zielerreichung). Aber auch Aspekte wie Intuitivität und Lernbarkeit lassen sich durch Verwendung unterschiedlicher Modalitäten steigern.

Demgegenüber steht der viel schlechter quantifizierbare – aber häufig viel direkter erfahrbare – Nutzen von Multimodalität auf der hedonischen Seite, z.B. bzgl. der Ästhetik und der wahrgenommenen „Persönlichkeit“ des Systems. So wirken multimodale Geräte häufig innovativer, äußerlich ansprechender, abwechslungsreicher, und sie versprechen dadurch mehr Spaß bei der Interaktion, sofern die verschiedenen Modalitäten adäquat funktionieren. Dies kann dazu führen, dass multimodale Systeme oft vom Nutzer präferiert werden, auch wenn die Interaktionsmöglichkeiten gar nicht verwendet werden.

Experimentelle Ergebnisse

In den vergangenen Monaten wurden an der TU Berlin verschiedene Experimente zum Einfluss unterschiedlicher Eingabemodalitäten auf die wahrgenommene Qualität durchgeführt. Sie zeigen, dass

- multimodale Systeme insbesondere bei hedonischen Qualitätsaspekten unimodalen überlegen sind, auch wenn unimodale eine höhere pragmatische Qualität erzielen (Untersuchung von *Attentive Displays* in [5])
- die Gesamtqualität und die Attraktivität eines multimodalen Systems aus den Qualitäten der Einzelmodalitäten als Linearkombination vorhergesagt werden kann, wobei die Gewichtungen der Einzelmodalitäten bei der Gesamtqualität etwa ihrer tatsächlichen Nutzungshäufigkeit entsprechen, während bei hedonischen Aspekten auch wenig genutzte Modalitäten eine Rolle spielen (multimodale *Voicebox* in [6]).

Offene Forschungsfragen betreffen insbesondere die Verwendungshäufigkeit der Modalitäten, ihr Einfluss auf Effektivität und Effizienz, sowie der Zusammenhang zwischen der Nutzung von Ein- und Ausgabemodalitäten. Mit diesem Wissen und neuen experimentellen Daten könnten dann bessere Vorhersagemodelle aufgestellt werden, die die Verwendung der Interaktionsmodalitäten und ihren Einfluss auf verschiedene Qualitätsaspekte besser vorhersagen und die Qualität multimodaler Systeme damit planbar machen.

Wir bedanken uns bei der DFG für die Förderung dieses Projektes (MO 1038/6-1).

Literatur

- [1] Möller, S., Engelbrecht, K.-P., Kühnel, C., Wechsung, I., Weiss, B.: Evaluation of multimodal interfaces for ambient intelligence, in: *Human-Centric Interfaces for Ambient Intelligence*, Elsevier, S. 347–370, 2010.
- [2] Coutaz, J., Nigay, L., Salber, D., Blandford, A.E., May, J., Young, R.M.: Four easy pieces for assessing the usability of multimodal interaction: The CARE properties, in: *Human-Computer Interaction, Proc. Interact'95*, Chapman & Hall, London, S. 115–120, 1995.
- [3] Bernsen NO, Dybkjær L: A theory of speech in multimodal systems, in: *Proc. ESCA Workshop on Interactive Dialogue in Multi-Modal Systems*, Europ. Speech Comm. Assoc., Kloster Irsee, S 105–108, 1999.
- [4] Wickens, C. D.: Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50, S. 449–455, 2008.
- [5] Wechsung, I., Engelbrecht, K.-P., Schaffer, S., Seebode, J., Metze, F., Möller, S.: Usability-Evaluation multimodaler Schnittstellen: Ist das Ganze die Summe seiner Teile? *Mensch und Computer 2009*, Oldenbourg, S. 495–498, 2009.
- [6] Wechsung, I., Engelbrecht, K.-P., Nauman, A., Schaffer, S., Seebode, J., Metze, F., Möller, S.: Predicting the Quality of Multimodal Systems Based on Judgements of Single Modalities, in: *Proc. Interspeech 2009*, S. 1827–1830, 2009.