

Multimodale Interaktion

Lennard Hamann
INF-M3 – Anwendungen 2 - Wintersemester 2007/2008
18. Dezember 2007

Agenda

- ▶ Einleitung
- ▶ SmartKom
- ▶ MATCH
- ▶ Ergebnisse

Motivation

- ▶ Starke Zunahme an IT-Anwendungen in allen Lebensbereichen
- ▶ Rasch zunehmende Komplexität von IT-Systemen
- ▶ Immer geringere Zeit, welche den Benutzern zum Ausführen von Aufgaben und dem Erlernen von Bedienkonzepten zur Verfügung steht

Ziele

- ▶ Intelligente Benutzerschnittstelle, die den natürlichen Kommunikationsstil von Computerlaien akzeptiert und so den Zugriff auf Informationen und Anwendungen erleichtert
- ▶ Für den Menschen intuitive und benutzergerechte Mensch-Technik-Interaktion

Ziele

- ▶ Weg von produktspezifischer Bedienung, bei der der Benutzer gezwungen ist, die "Sprache" der Maschine zu lernen
- ▶ Dem Benutzer soll nicht bewußt sein, daß er mit einer Maschine interagiert

Multimodale Interaktion

- ▶ **Modalität**
 - ▶ Sinnesorgan, mit dem der Benutzer die Ausgaben vom Computer wahrnimmt
 - ▶ Sensor, mit dem ein Computer die Eingaben vom Benutzer empfangen kann
- ▶ **Multimodale Interaktion/Symmetrische Multimodalität**
 - ▶ Sprache, Gesten, Mimik etc. zur Ein- **UND** Ausgabe

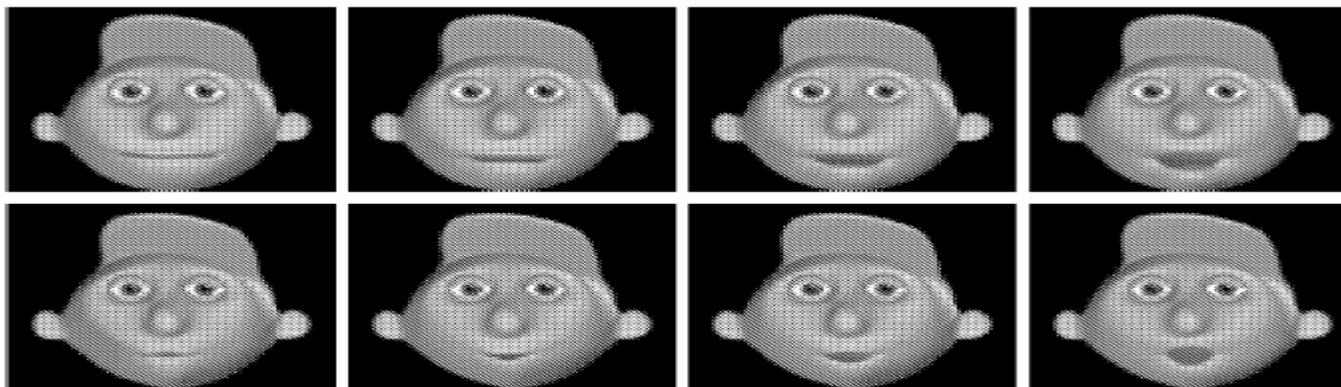


Abbildung 1:
Visime in
SmartKom
(aus [Wahlster
2002b])

SmartKom

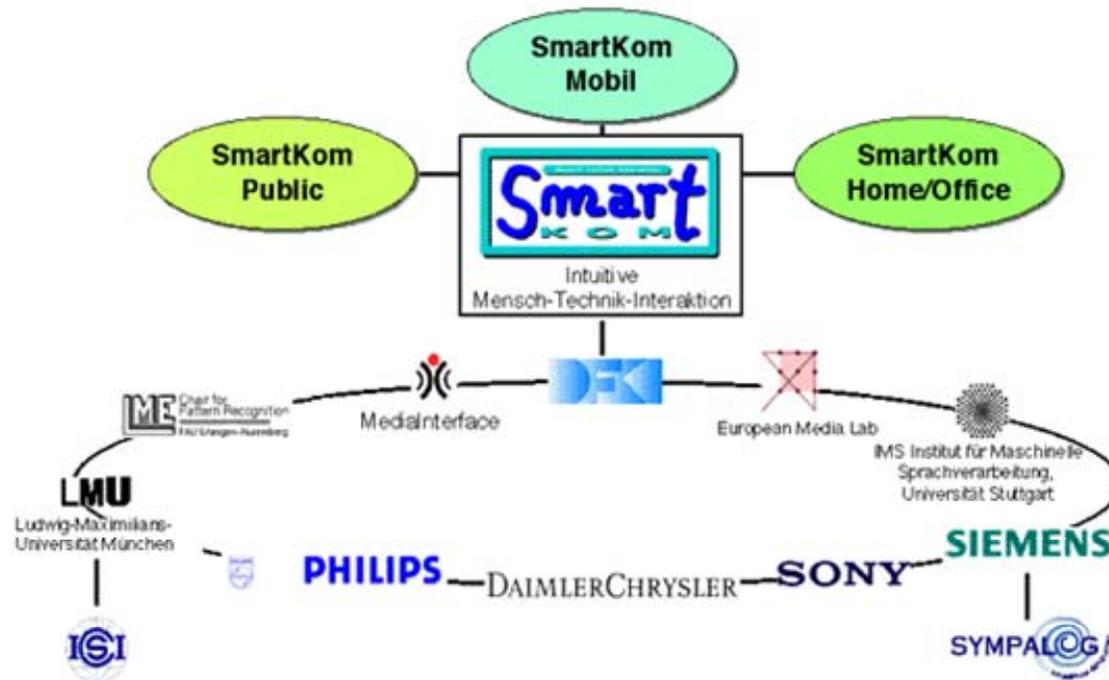


Abbildung 1: Das SmartKom-Konsortium

SmartKom

- ▶ Dialogische Mensch-Technik-Interaktion
- ▶ Benutzer delegiert Aufgaben an den Präsentationsagenten
- ▶ Sowohl der Benutzer als auch das System starten Interaktionen



Abbildung 1:
Smartkaus (aus
[Wahlster 2002b])

Dialogische Mensch-Technik-Interaktion

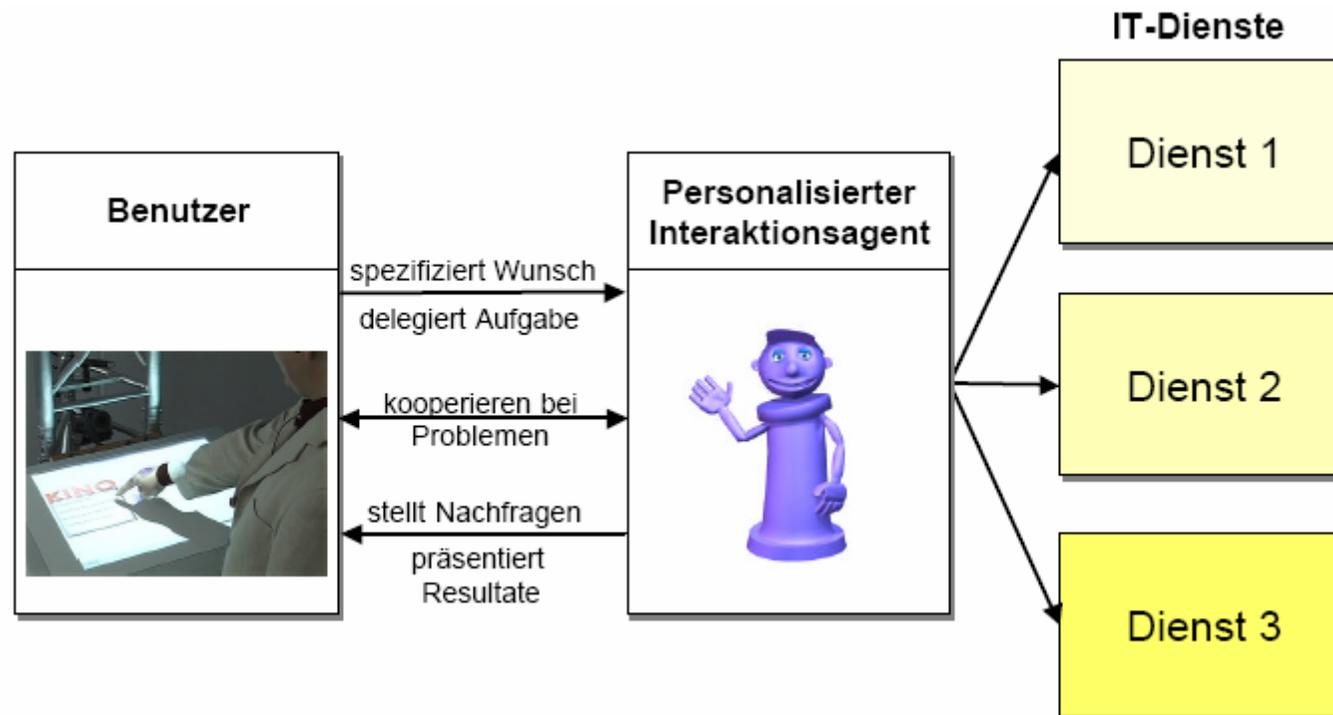


Abbildung 1: Situated Delegation-oriented Dialogue Paradigm (aus [Wahlster 2002b])

SmartKom: Paradigmen

- ▶ Radikaler Bruch mit dem traditionellen Desktop, da die Bürometapher für Mobile Geräte, Ambient Living etc. nicht angemessen ist.
- ▶ Abkehr vom WIMP-Prinzip, (Windows, Icons, Mouse Pointer), da die Fenstertechnik auf Mobilgeräten problematisch ist und der Mausklick durch komplexe Zeigegeesten ersetzt wird.

SmartKom: Paradigmen

- ▶ Keine Kommandos an den Kommunikationsassistenten, sondern Dialog mit gemischter Initiative: Ziel ist die Aufgabendelegation, ihre Spezifikation sowie notwendige Hilfestellungen bei der Problemlösung.
- ▶ Zielt auf situierte Dialoge, da der personalisierte Interaktionsagent immer in bestimmten Umgebungen auftritt (z.B. Landkarte, Sitzplan eines Kinos, Fernsehprogramm).

Animierter Präsentationsagent

- ▶ Lenkt die Aufmerksamkeit des Benutzers
- ▶ Stellt Querverweise zwischen graphischen und textuellen Elementen durch beidhändiges Zeigen her
- ▶ Hilft dem Benutzer (z.B. Umgang mit Formular)



Abbildungen 1 und 2: Präsentationsagent Smartakus aus SmartKom (aus [Wahlster 2003])

Anwendungsfall: Kinokarten bestellen

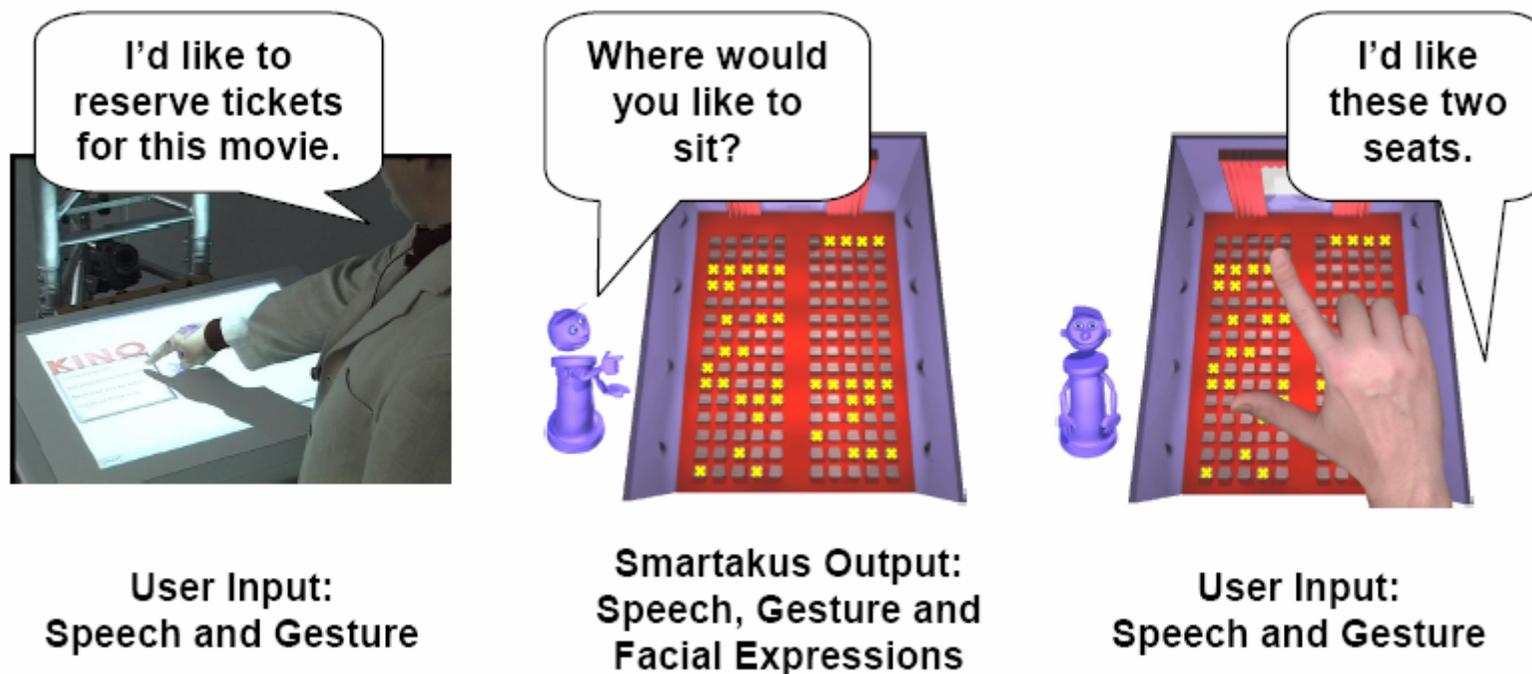


Abbildung 1: Multimodale Interaktion mit SmartKom (aus [Wahlster 2002b])

Interpretation von mimischen Gesten

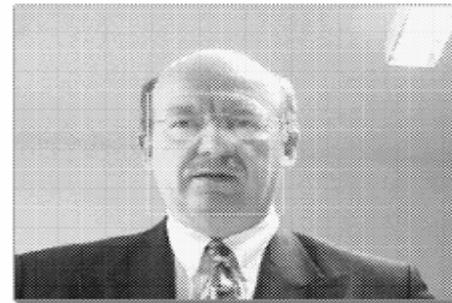
(1) Smartakus: Here you see the CNN program for tonight.

(2) User: That's great.



(3) Smartakus: I'll show you the program of another channel for tonight.

(2') User: That's great.



(3') Smartakus: Which of these features do you want to see?

Abbildung 1: Interpretation von mimischen Gesten (aus [Wahlster 2002b])

SmartKom: Modulare Architektur

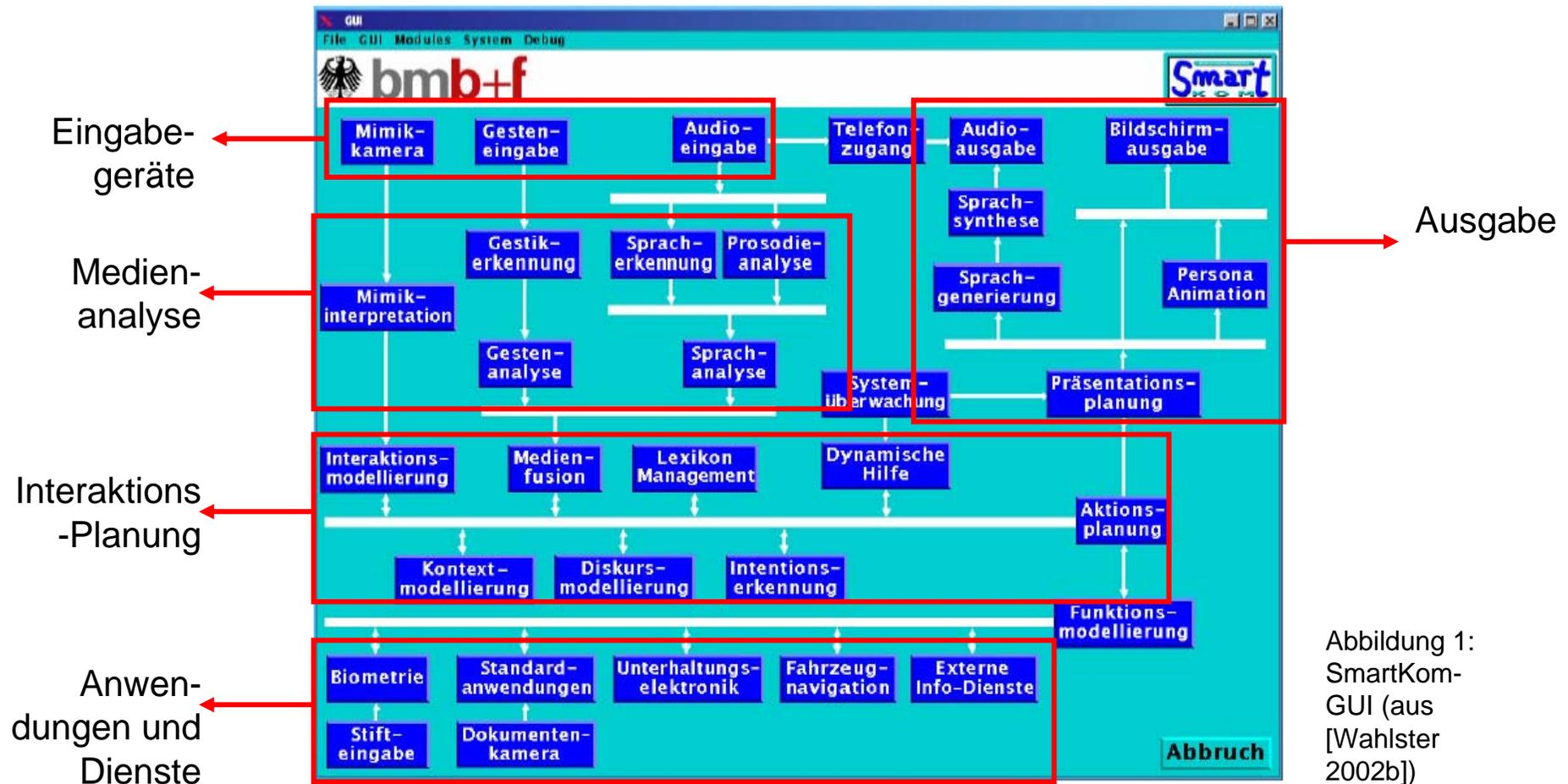


Abbildung 1:
SmartKom-
GUI (aus
[Wahlster
2002b])

Multimodal Access To City Help (MATCH)

- ▶ AT&T Projekt
- ▶ Restaurant und U-Bahn Informationen abrufen
- ▶ Eingabe mit Sprache und Stift
- ▶ Berücksichtigt außerdem den Dialog-Kontext und Benutzerpräferenzen

Anwendungsfall: Restaurant finden

- ▶ Optionen bei der Eingabe:
 - ▶ Sprache: 'Show cheap italian restaurants in chelsea'
 - ▶ Stift und Sprache: 'Show cheap italian restaurants in this neighborhood' + Zeigen
 - ▶ Stift:

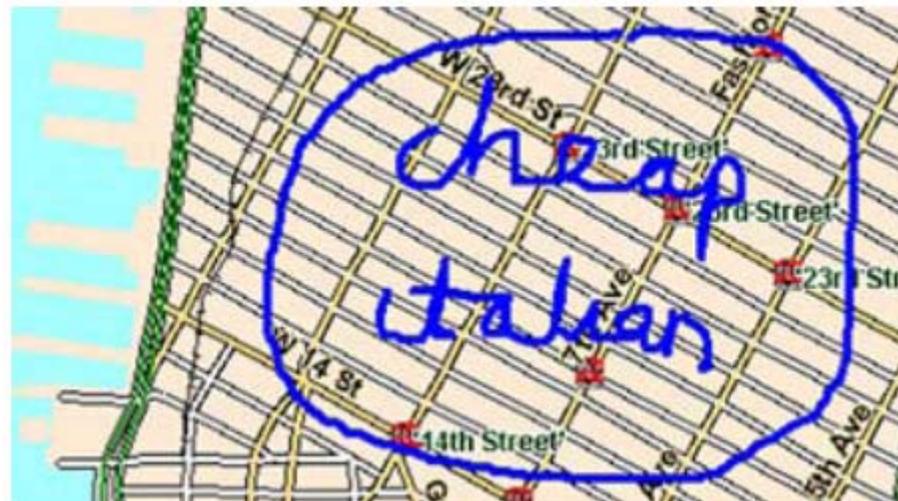


Abbildung 1: Unimodale eingabe per Stift (aus [Johnston u. a. 2002])

Ergebnisse

- ▶ Benutzer wechseln zwischen unimodaler und multimodaler Eingabe in Abhängigkeit von der auszuführenden Aktion
- ▶ Multimodale Eingaben geschehen nicht synchron, weder in der Mensch-Technik Interaktion noch in der Mensch-Mensch Interaktion

Ergebnisse

- ▶ Ausschließlich auf Speak-and-point ausgerichtete Systeme sind nicht ausreichend
 - ▶ Nur 20% der Gesten in der Mensch-Mensch Interaktion sind Zeigegesten
- ▶ Sprache ist nicht der einzige oder wichtigste Informationsträger in der Interaktion
- ▶ Multimodale Sprache ist einfacher als natürliche oder unimodale Sprache

Vielen Dank für die Aufmerksamkeit

Literatur

- ▶ [Alexandersson 2002] Alexandersson J. : Overlay - a non-monotonic operation for discourse processing. In: Loria (Laboratoire lorrain de recherche en informatique et ses applications), Nancy, 2002.
- ▶ [Cohen u. a. 1997] Cohen P.R. ; Johnston, M. ; McGee, D. ; Oviatt, S. ; Pittman, J. ; Smith, S. ; Chen, L. ; Clow, J. : QuickSet: multimodal interaction for simulation set-up and control. In: Proceedings of the fifth conference on Applied natural language processing, Washington, DC, 1997, pp. 20 – 24.
- ▶ [Johnston u. a. 2002] Johnston, M. ; Bangalore, S. ; Vasireddy, G. ; Stent, A. ; Ehlen, P. ; Walker, M. Whittaker, S. ; Maloor, P. : MATCH: An Architecture for Multimodal Dialogue Systems. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, July 2002, pp. 376-383.

Literatur

- ▶ [Oviatt 1997a] Oviatt, S.L. : Multimodal interactive maps: Designing for human performance. *Human-Computer Interaction* 12, (1997), pp. 93–129.
- ▶ [Oviatt u.a. 1997b] Oviatt, S.L. ; DeAngeli, A. ; Kuhn, K. : Integration and synchronization of input modes during multimodal human-computer interaction. In *Proceedings of Conference on Human Factors in Computing Systems CHI'97* (March 22–27, Atlanta, GA). ACM Press, NY, 1997, pp. 415–422.
- ▶ [McNeill 1992] McNeill, D. : *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press, Chicago, IL, 1992.

Literatur

- ▶ [Wahlster 2002a] Wahlster W. : Multimodale Interaktion und Interface Agenten: Trends für Morgen und Übermorgen. In: USEWARE, Darmstadt, 2002.
- ▶ [Wahlster 2002b] Wahlster W. : SmartKom: Fusion and Fission of Speech, Gestures, and Facial Expressions. In: Proceedings of the 1st International Workshop on Man-Machine Symbiotic Systems, Kyoto (Japan), 25.-26.11.2002, pp. 213-225.
- ▶ [Wahlster 2003] Wahlster W. : SmartKom: Symmetric Multimodality in an Adaptive and Reusable Dialogue Shell. In: Krahl, R., Günther, D. (eds): Proceedings of the Human Computer Interaction Status Conference 2003 03.06.2003 Berlin: DLR, pp. 47-62.