

Experimentelle Untersuchung von Spracheingaben unter kognitiver Belastung zur Benutzermodellbildung

Barbara Großmann-Hutter, Christian Müller *
Fachbereich Informatik, Universität des Saarlandes
Postfach 151150, D-66041 Saarbrücken
{ barbara | cmueller}@cs.uni-sb.de

16. Juli 1999

Zusammenfassung

Hier wird ein Experiment vorgestellt, in welchem Benutzer unter variierender kognitiver Belastung natürlichsprachliche Eingaben an ein System machen. Unser Hauptinteresse liegt dabei in der Erhebung genauer quantitativer Messergebnisse von linguistischen Merkmalen der Eingaben und Antwort- und Aktionszeiten. Die so gewonnene Datensammlung wird zum Lernen eines Bayes'schen Netzes verwendet, mit dessen Hilfe, nach Abschluss des Lernvorganges, Systemanpassungen an aktuelle Umweltsituationen eines Benutzers ermöglicht werden.

1 Einleitung

Systeme, die sich auf ihre Benutzer einstellen, können das in zweierlei Hinsicht tun: zum einen können sie sich an individuelle Unterschiede einzelner Benutzer anpassen und zum anderen an verschiedene Umweltsituationen, in denen sich Benutzer befinden. Wenn man berücksichtigt, dass Computer heute mehr und mehr Einzug in unser ganz alltägliches Leben halten, wird klar, dass damit eine Erhöhung des Einflusses der Umwelt einhergeht, der die Kommunikation zwischen System und Benutzer mehr und mehr beeinträchtigen kann. Deshalb gilt unser derzeitiges Interesse solchen situativen Einflüssen. Wir untersuchen linguistische Merkmale von Benutzereingaben, um daraus Rückschlüsse auf die aktuelle Situation, in der sich der Benutzer gerade befindet, ziehen zu können.

1.1 Szenario

Im Dialogsystem READY wurden Erfahrungen zur Anpassung von Dialogen an die Ressourcenbeschränkungen eines Benutzer gesammelt. Im Projekt REAL (Wahlster, Blocher, Baus,

*Dieses Projekt wird im Rahmen des Sonderforschungsbereiches "Ressourcenadaptive kognitive Prozesse", SFB 378, Projekt B2, READY von der Deutschen Forschungsgemeinschaft gefördert

Stopp & Speiser, 1998) wurden anytime-Algorithmen zur Raumbeschreibung entwickelt, die eine Anpassung an Ressourcenbeschränkungen eines Benutzer erlauben. Darauf aufbauend soll nun ein Assistenz-System zur Navigationshilfe in einem Leitszenario "Flughafen" entwickelt werden. Hierzu untersuchen wir Möglichkeiten, kognitiv belasteten Passagieren Wegbeschreibungen, Auskünfte u.ä. in adäquater Form anzubieten.

Die Ein- und Ausgaben der Navigationshilfe sollen mittels unterschiedlicher Medien erfolgen können. Situationsangepasst soll die Kommunikation sowohl über Display, als auch über sprachliche Interaktion möglich sein. Damit kann sowohl der erfahrene Benutzer mit wenigen Tastendrücken an die benötigte Information kommen, als auch der mit Gepäck beladene, der keine Hand mehr frei hat, oder der Eilige, der seine Augen beim Laufen offen halten muss, per Sprache mit dem System kommunizieren.

Der sprachlichen Interaktion mit dem System gilt nun unser Augenmerk.

1.2 Ziel des Experimentes

Um eine Adaption des Navigationshilfesystems an den Benutzer und seine augenblickliche Umweltsituation zu ermöglichen, müssen zunächst Möglichkeiten gefunden werden, wie über natürlichsprachliche Eingaben eines Benutzers Rückschlüsse auf dessen kognitive Belastung gezogen werden können .

Die Aspekte der prinzipiellen Erkennbarkeit kognitiver Belastung in sprachlichen Äusserungen wurde bereits in (Berthold, 1998) aufgezeigt und ein Benutzermodell zur Darstellung der Zusammenhänge verschiedener beobachtbarer Symptome in (Berthold & Jameson, 1999) vorgestellt, siehe Abb. 1.

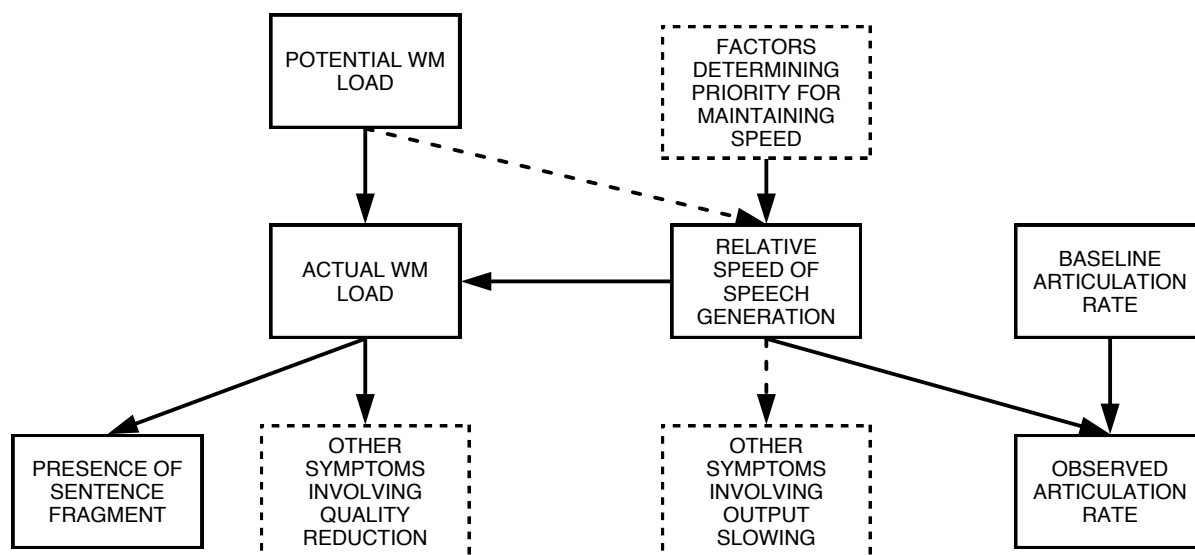


Abbildung 1: Einfaches Bayes'sches Netz zur Interpretation kognitiven Belastung bei Spracheingaben

Aufbauend auf diesen Arbeiten wollen wir nun quantitative Daten zum Zeit- und Qualitätsverhalten von Benutzeräusserungen unter kognitive Belastung sammeln. Unser Hauptinteresse liegt dabei in der Frage, ob ein Trade-Off zwischen Zeit und Qualität bei Benutzeräusserungen

existiert und wie er aussieht. Wir vermuten, dass mit zunehmender Äusserungsgeschwindigkeit die Äusserungsqualität abnimmt, und umgekehrt. Der Experimentaufbau und -ablauf soll sowohl diesem Ziel Rechnung tragen, als auch das Szenario ausreichend repräsentieren.

2 Entwurf des Experimentes

Die Gestaltung der Experimentalumgebung (siehe Abb.3) ist eng an den zukünftigen Einsatzbereich "Flughafen" angelehnt: während sich der Benutzer durch den Flughafen bewegt, muss er sowohl bestimmte Aufgaben wie z.B. das "Einchecken" erledigen, als auch das Endziel zum "Gate xx" zu gelangen bis zu einem festgelegten Zeitpunkt erreicht haben. Ein Hilfesystem soll den Benutzer dabei unterstützen.

Dieses Szenario spiegelt sich in den Aufgaben der Versuchspersonen wider: In dem Experiment sind sowohl Fragen an das fiktive Hilfesystem zu stellen, als auch eine Figur auf einem Lageplan des Flughafens zu bestimmten Orten zu führen. Wir haben die Aufgaben so gewählt, dass sich die bei der peripheren Ausführung in Anspruch genommenen Subsysteme des Arbeitsgedächtnisses nicht überschneiden (Sprachgenerierung, motorisches Figurbewegen), es aber sehr wohl zu Konkurrenzsituationen im Bereich der Zentralen Exekutive kommen kann (Memorieren des Endzieles, Bild-Sprach-Umsetzung), was speziell unter Zusatzanforderungen wie Zeitdruck zu Engpässen führen muss.

2.1 Sprachaufgabe

Die Primäraufgabe der Versuchspersonen wird es sein, Anfragen an ein Assistenzsystem zu artikulieren, wie es im Rahmen des Flughafenszenarios entworfen werden soll. Um den größtmöglichen Realismus zu erreichen, wurde im Vorfeld eine Feldstudie am Frankfurter Flughafen durchgeführt. Dabei wurde ein Katalog von Anfragen zusammengestellt, wie sie an einem konventionellen Informationsstand anfallen.

Um eine freie Formulierung von Seiten der Versuchspersonen zu ermöglichen, werden die Konzepte, die artikuliert werden sollen, nicht sprachlich, sondern in Form von Bildern präsentiert (vgl. Abbildung 2). Die Bilder bestehen aus mehreren einfachen Komponenten, die mit einer eindeutigen Syntax verbunden sind. Die Einzelkomponenten werden den Versuchspersonen vor dem Experiment vorgestellt, jedoch nie in dem Zusammenhang, in dem sie später präsentiert werden. Dadurch wird gewährleistet, dass die Semantik des Bildes erkannt wird, ohne dass die Anfrage bereits vorformuliert werden konnte. Die Bilder unterscheiden sich in ihrer Komplexität. Abbildung 2 a) drückt das eher einfache Konzept aus, dass sich der Reisende auf der Suche nach einer Apotheke befindet. Wir erwarten in diesem Fall z.B. die Äusserung „Wo ist eine Apotheke?“ oder „Gibt es hier eine Apotheke?“. Abbildung 2 b) ist bereits etwas komplexer. Es kommt das Konzept zum Ausdruck, dass der Reisende mit dem Flughafen-Shuttle von Terminal 1 nach Terminal 2 fahren will. Da die Einzelkomponenten der Bilder und die Einrichtungen am Flughafen Frankfurt vor Beginn des Experimentes vorgestellt werden, erwarten wir in diesem Fall Äusserungen wie „Wo befindet sich die Abfahrtstelle des Shuttles, das zum Terminal 2 fährt?“. Abbildung 2 c) ist noch komplexer und repräsentiert ein Konzept, das beispielsweise wie folgt artikuliert werden könnte: „Ist es erlaubt, ein Taschenmesser im Bord-Gepäck mit sich zu führen?“.

Um die Komplexität als Störvariable auszuschließen, werden die Bilder anhand der Anzahl der Einzelkomponenten in drei Kategorien eingeteilt. Jede der experimentellen Bedingungen wird die gleiche Anzahl Bilder einer Komplexitätskategorie enthalten.

Die Sätze, die die Versuchspersonen erzeugen, werden per Mikrophon aufgezeichnet, und anschließend auf Symptome der kognitiven Belastung hin untersucht.

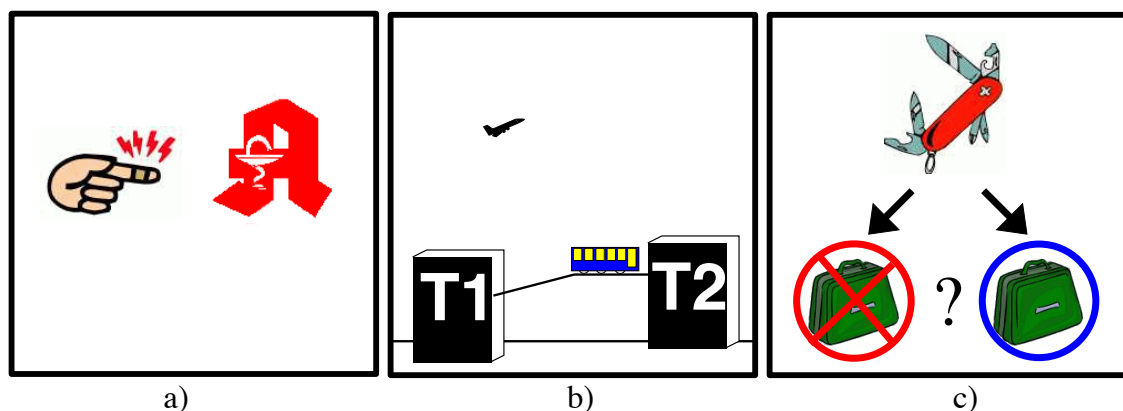


Abbildung 2: a) „Wo ist die Apotheke?“ b) „Wie komme ich mit dem Shuttle-Bus zum Terminal 2?“ c) „Darf ich ein Taschenmesser im Handgepäck mitführen?“

2.2 Navigationsaufgabe

Ziel dieser Aufgabe ist es, eine ständige kognitive Grundbelastung der Versuchsperson zu gewährleisten. Hierzu wird ihr der Ausschnitt eines Flughafenplanes präsentiert, auf dem sowohl Endziele (6 blaue Rechtecke), als auch diverse Gebäude mit Eingängen (gelbe Quadrate als Teilziele) und Weghindernisse dargestellt sind, siehe Abb. 3.

Der Ablauf gestaltet sich wie folgt: Zunächst bekommt die Versuchsperson kurz die Nummer, z.B. „5656“, eines Endzieles präsentiert, das sie sich einprägen soll. Anschliessend wird auf dem Plan das 1. Teilziel, das von der Figur erreicht werden soll, durch Aufleuchten des Gebäudeeinganges kurz angezeigt. Die Versuchsperson bewegt nun die Figur mittels der Pfeiltasten der Tastatur auf den Wegen des Planes zum Teilziel. Nach Betreten eines Teilzieles wird jeweils das nächste Teilziel präsentiert, bis 5 Teilziele abgearbeitet worden sind. Die 6 blauen Zielrechtecke werden nach Erreichen dieses letzten Teilzieles mit Variationen der Zielnummer beschriftet z.B. „5656“, „5566“.... Nun muss die Versuchsperson sich an die anfangs präsentierte Zielnummer erinnern und mit der Figur das entsprechende Ziel anlaufen. Damit gilt die Aufgabe als abgeschlossen.

Um Interferenzen zwischen den Aufgaben auszuschließen, werden 3-4 unterschiedliche Pläne verwendet, die von Aufgabe zu Aufgabe wechseln. Die Komplexität dieser Pläne wird gleichwertig gehalten.

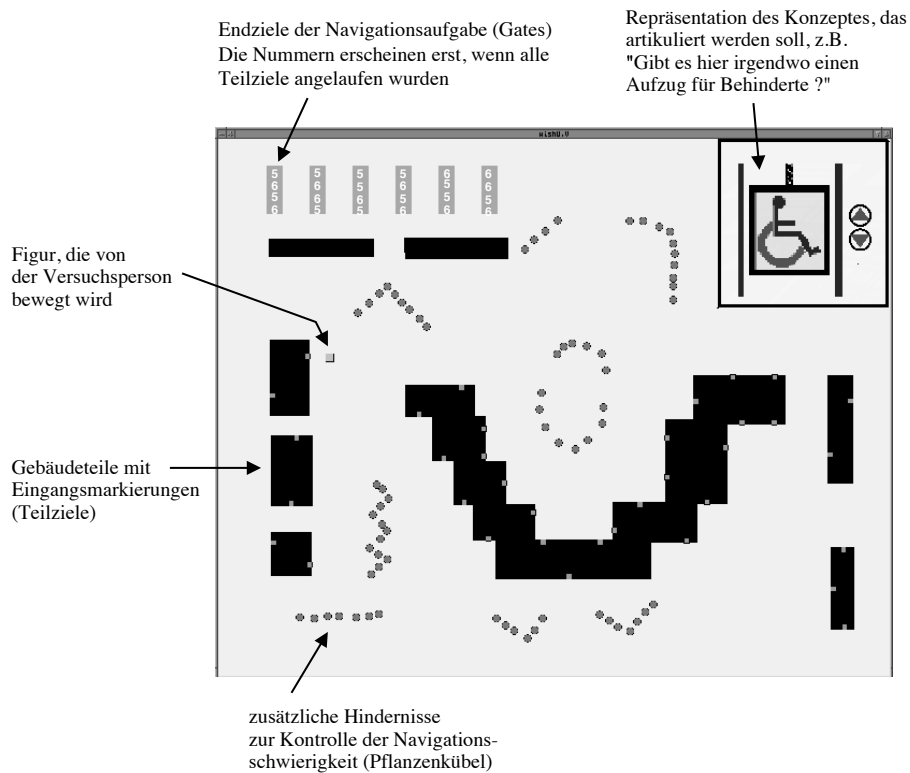


Abbildung 3: Navigationphase mit Bildanzeige

2.3 Design

Die Hauptexperimentphase gliedert sich in die in Tabelle 1 aufgeführten Zellen. Zunächst bearbeitet eine Versuchsperson Aufgaben aus Zellen ohne zusätzliche Navigationsaufgabe. Die aufgezeichneten Daten werden sowohl zur Ermittlung einer Baseline der Artikulationsgeschwindigkeit einer Versuchsperson benötigt, als auch für Vergleiche mit Messungen aus den Doppelaufgabenzellen, die sich im 2. Teil des Hauptexperimentes anschließen. Hier wird in 4 Blöcken zu jeweils 3 Aufgaben abwechselnd das Gewicht auf Zeitdruck oder auf Qualität gelegt und die Versuchspersonen entsprechend instruiert.

	Vorrang	
	Zeit	Qualität
ohne Navigieren		
mit Navigieren		

Tabelle 1: Variation der unabhängigen Variablen in den Versuchszellen

In den Phasen mit paralleler Aufgabenausführung soll die Versuchsperson möglichst konstant mit Navigieren beschäftigt sein, während sie Anfragen an das Hilfesystem artikuliert. Die beiden Aufgaben werden gekoppelt, indem kurz nachdem sich die Figur auf den Weg zum Teilziel befindet ein Bild angezeigt wird, das den aktuellen Frageinhalt darstellt. Nun muss die Versuchsperson die Frage artikulieren, während gleichzeitig die Figur ohne längere Pausen (bis 2sec) bis

zum nächsten Teilziel zu bewegen ist.

Die Wege zwischen den einzelnen Teilzielen sind derart gestaltet, dass die Zeit in jedem Fall ausreicht um eine Frage zu artikulieren.

Die einzelnen Phasen und Aktionen des Experimentes werden so aufgezeichnet, dass sich zum einen gegebenenfalls die Verzahnung der beiden parallelen Aufgaben aufzeigen lässt und zum anderen für jede Aufgabe Zeit- und Qualitätsangaben ermittelt werden können.

2.4 Zu untersuchende Symptome

(Berthold, 1998) untersuchte eine Reihe von Arbeiten zur Sprachproduktion und kognitiver Belastung. Er stellte eine Liste von Merkmalen auf, die als Belastungssymptome interpretiert werden können. Die Wichtigsten werden in Tabelle 2 aufgeführt und in dem darauffolgenden Abschnitt kurz erläutert. Die Symptome in der linken Spalte betreffen die Äusserungsqualität, die Symptome in der rechten Spalte die Äusserungsgeschwindigkeit.

Wir vermuten, dass es einen Trade-Off zwischen Qualität und Geschwindigkeit gibt und erwarten folgende Beobachtungen:

- In der Bedingung, in der die Versuchspersonen angehalten werden, auf die Qualität der Äusserung zu achten, sollten verstärkt die Geschwindigkeitsymptome auftreten.
- In der Zeitdruck-Bedingung dagegen sollten die Qualitätssymptome überwiegen.
- Insgesamt sollten in den Bedingungen mit Nebenaufgabe mehr Symptome auftreten als in den unbelasteten Bedingungen.

Qualitätssymptome		Geschwindigkeitssymptome	
Merkmal (Anzahl)	Tendenz*	Merkmal	Tendenz
Satzfragmente	+	Artikulationsgeschwindigkeit	–
Fehlansätze	+	Sprechgeschwindigkeit	–
Syntaktische Strukturfehler	+	Einsatzlatenz (Dauer)	+
		Stille Pausen (Anzahl)	+
		Stille Pausen (Dauer)	+
		Gefüllte Pausen (Anzahl)	+
		Gefüllte Pausen (Dauer)	+
		Wiederholungen	+

Tabelle 2: Symptome der kognitiven Belastung. * „+“ bedeutet, dass mit der kognitiven Belastung das Maß dieses Symptoms zunimmt; „–“ bedeutet, dass es abnimmt.

Die Definition der Symptome wird wie folgt angenommen: *Satzfragmente* sind Äusserungen, die während der Artikulation abgebrochen, und nicht vervollständigt werden (siehe Bsp.1 unten). Ein *Fehlansatz* ist ein Spezialfall eines Satzfragmentes, bei dem in der nächsten Äusserung eine Korrektur erfolgt (2). Durch einen *Syntaktischen Strukturfehler* wird eine Äusserung

ungrammatisch (3). *Pausen* sind Phasen während der Äusserung, in denen nicht artikuliert wird (*Stille Pausen*), oder die Artikulation inhaltslos ist (*Gefüllte Pausen*). Die *Sprechgeschwindigkeit* definiert sich als die Anzahl geäusselter Spracheinheiten pro Zeiteinheit. Die *Artikulationsgeschwindigkeit* (im engeren Sinne) ist die Sprechgeschwindigkeit abzüglich aller gefüllter und stiller Pausen. Wird ein Äusserungssegment beliebiger Länge unverändert reproduziert, spricht man von einer *Wiederholung* (4). Die Beispielsätze stammen aus einer Feldstudie von (Berthold, 1998), bei der die Sprache von Personen untersucht wurde, die nach einer Autopanne mit einem Reparaturservice telefonieren.

- (1) Die Leitung, warum hört denn die ... (auf zu funktionieren)
- (2) In der Mitte ist ein großes rundes Ding, so'n...Gucken Sie mal, ob sie links den Ölmesstab sehen.
- (3) Also nicht nur ich rausfinden, sondern Du musst ihn auch beheben
- (4) Ich hab' noch < PAUSE > ich hab' noch 'ne Frage überhaupt.

3 Verwertung der Ergebnisse

Wie können wir nun unsere Experimentergebnisse so umsetzen, dass unser Hilfesystem entsprechende Entscheidungen zur Adaption an eine aktuelle Benutzersituation treffen kann ?

In einem ersten Experiment haben wir die Ausführung von unterschiedlich präsentierten Anweisungsfolgen bei kognitiv belasteten Benutzern untersucht (March, 1999). Anschliessend haben wir die Experimentdaten zum Lernen eines Bayes'sches Netzes mit vorgegebener Struktur benutzt (Großmann-Hutter, Jameson & Wittig, 1999) und waren so in der Lage, auf Grund einiger bekannter Symptome mit mehr oder weniger grosser Sicherheit ein Gesamtbild der Situation, in der sich der Benutzer befindet, zu entwickeln. Diese Situationsbeschreibung kann dann zur Benutzeradaption genutzt werden.

Zur Aufarbeitung von Sprachsymptomen mit Bayes'sches Netzen wurde das in Abb. 1 dargestellte Teilnetz entworfen, eingesetzt, und aufgezeigt, welche Systemanpassungen an einen Benutzer damit ermöglicht werden können (Berthold & Jameson, 1999).

Aufbauend auf den Ergebnissen dieser Arbeiten werden wir die Daten dieses Experimentes zum Lernen eines weiteren Bayes'sches Teilnetzes verwenden und so für Entscheidungsprozesse des Systems nutzbar machen. Hierzu arbeiten wir auch an Techniken zur optimalen Ausnutzung der Experimentdaten zum Lernen und an Möglichkeiten des Aufbaus von Bayes'schen Netzen aus u.U. gelernten Teilnetzen (Wittig, 1999; Jameson, Großmann-Hutter, March & Rummer, 1999). Ein grosser Vorteil in der Verwendung solcher Bayes'sches Netze zur Benutzermodellierung liegt in der Möglichkeit, einmal erkannte Umweltsituationen in verschiedene Bereiche des Hilfesystems einfließen zu lassen und so möglichst umfassende Anpassungen an den Benutzer vornehmen zu können.

4 Ausblick

Im Augenblick sind wir dabei den Entwurf umzusetzen und die Experimentalumgebung aufzubauen. Wir sind zuversichtlich bis zum Workshop Ergebnisse der ersten Experimente präsentieren zu können. Den ABIS-Teilnehmern wird die Gelegenheit geboten, probeweise einen Teil des Experimentes durchzuführen.

References

- Berthold, A. (1998). *Repräsentation und Verarbeitung sprachlicher Indikatoren für kognitive Ressourcenbeschränkungen*. Diplomarbeit, Universität des Saarlandes.
- Berthold, A. & Jameson, A. (1999). Interpreting symptoms of cognitive load in speech input. In J. Kay (Hrsg.), *User modeling: Proceedings of the Seventh International Conference, UM99*. Wien, New York: Springer Wien New York.
- Großmann-Hutter, B., Jameson, A. & Wittig, F. (1999). Learning Bayesian networks with hidden variables for user modeling. In *Proceedings of the IJCAI99 Workshop „Learning About Users“*. Stockholm, July 31st 1999.
- Jameson, A., Großmann-Hutter, B., March, L. & Rummer, R. (1999). *Creating an empirical basis for adaptation decisions*. (Manuscript submitted for publication.)
- March, L. (1999). *Ressourcenadaptive Instuktionen in einem Hotline-Szenario*. Diplomarbeit, Universität des Saarlandes.
- Wahlster, W., Blocher, A., Baus, J., Stopp, E. & Speiser, H. (1998). Ressourcenadaptierende Objektlokalisierung: Sprachliche Raumbeschreibung unter Zeitdruck. *Kognitionswissenschaft*, 7(3), 111-117.
- Wittig, F. (1999). Learning Bayesian networks with hidden variables for user modeling. In J. Kay (Hrsg.), *User modeling: Proceedings of the Seventh International Conference, UM99*. Wien, New York: Springer Wien New York. (Doctoral Consortium summary)