EINSATZ VON SPRACHANWENDUNGEN IM AMBIENT ASSITED LIVING (AAL) – RÜCKBLICK, AKTUELLER STAND UND ZUKUNFTSPERSPEKTIVE

Maik Hampicke

Fraunhofer Institut für Zuverlässigkeit und Mikrointegration (IZM), Berlin
Applikationszentrum ‚Smart System Integration'
maik.hampicke@izm.fraunhofer.de

Kurzfassung: Der Beitrag gibt einen Überblick über Anwendungen der Sprachverarbeitung, um Menschen in Ihrer selbständigen Lebensführung zu unterstützen. Im Vordergrund steht dabei die Zielgruppe der älteren Menschen, die von den Sprachanwendungen besonders profitieren können. Die Anwendungsbereiche für Sprachsteuerungen in diesem Bereich sind außerordentlich vielfältig. Neben Steuerungen von Funktionen im Smart-Home-Bereich können auch Kommunikationsgeräte, Rollstühle, Türöffner aber auch Haushaltgeräte und medizinische Geräte (Blutdruckmesser, Blutzuckermessgeräte etc.) mit natürlicher Sprache benutzt werden.

Es erfolgt ein Rückblick auf Anwendungen der Sprachverarbeitung und deren Nutzen, wie sie im DFG-Projekt SENTHA (1996 – 2003) ausführlich untersucht wurden. Beleucht werden dabei Faktoren, die die Benutzung und damit die Akzeptanz von Sprach- und Sprechererkennung beeinflussen.
Dem Einsatz von leistungsfähigen, integrierten (embedded) Sprachsteuerungen wird in Zukunft eine größere Bedeutung zukommen. Allerdings spielen damit weitere Faktoren der Systemintegration, der Aufbau und Verbindungstechnik und des Packaging eine größere Rolle, damit die Systeme in der Praxis verfügbar gemacht werden können.

1 Ambient Assisted Living

Die absehbare demographische Entwicklung der nächsten Jahrzehnte führt nicht nur in Deutschland zu einer veränderten Altersstruktur der Bevölkerung. Kinder, die heute geboren werden, leben im Schnitt 30 Jahre länger, als dies noch vor 100 Jahren der Fall war. Zugleich werden durch eine dauerhaft sinkende Geburtenrate den immer älter werdenden Menschen immer weniger Jüngere gegenüber stehen. Die daraus resultierende Überalterung der Bevölkerung stellt die sozialen Systeme vor außergewöhnliche Herausforderungen. So erfordert der erhöhte Pflegebedarf höheren personellen und finanziellen Aufwand, um die Lebensqualität der älteren Menschen zu erhalten. Dieser wird bei der zu erwartenden Entwicklung nicht zur Verfügung stehen. Eine Entlastung versprechen jedoch zahlreiche technische Unterstützungen, die die Selbständigkeit der Älteren auch erhalten bzw. verlängern, wenn unter normalen Bedingungen eine eigenständige Lebensführung nicht mehr möglich ist. Gerade die sensorische, mentale und mobile Leistungsfähigkeit lässt naturgemäß mit dem Alter nach, so dass hier Unterstützungssysteme erforderlich sind, die durch moderne Technologien geleistet werden können. In diesem Zusammenhang spricht man auch von „Ambient Assisted Living" (AAL) oder „Altersgerechten Assistenzsystemen für ein gesundes und unabhängiges Leben", um die Lebensqualität des Menschen in allen Lebensabschnitten durch neue Technologien zu erhöhen [1].
Technische Unterstützungen bieten den Älteren schon in vielen Bereichen eine Unterstützung bei der selbständigen Lebensführung an. Eine wichtige Rolle spielt hierbei der viel diskutierte „Smart-Home-Bereich" (vgl. [2], [3], [4] und [5]). Zu erwähnen sind hierbei u.a. Geräte zur

Messung von Vitalparametern (Blutdruck, Puls, Temperatur, EKG, etc.), verschiedene Assistenzsysteme, aber auch immer ausgefeiltere Haushalts- und Kommunikationsgeräte.

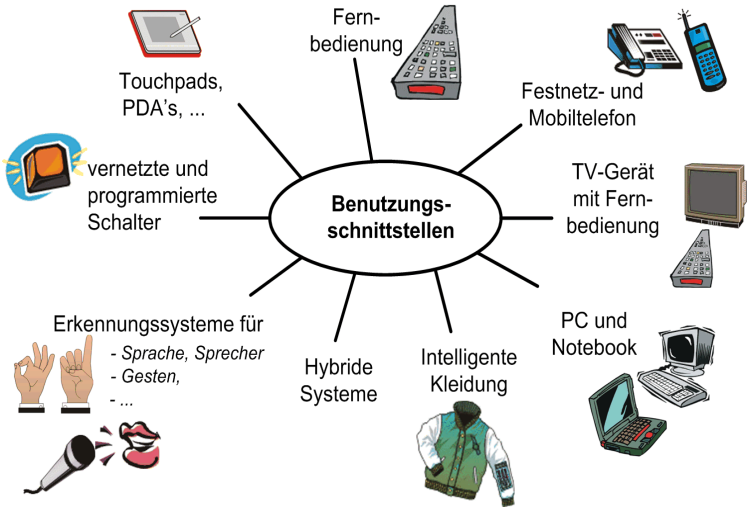

Abbildung 1 - Auswahl an Benutzungsschnittstellen

Allerdings können technische Lösungen den älteren Menschen nur dann unterstützen, wenn der Zugriff auf diese Funktionen ermöglicht werden kann. Vollkommen automatisierte Lösungen werden durch die Älteren weitgehend abgelehnt. Sprachsteuerungen nehmen jedoch eine Schlüsselstellung ein, weil die Sprache als natürliche Kommunikationsform verwendet wird.

2 Anwendungen der Sprachsteuerung

Die elektronische Sprachverarbeitung und insbesondere die Sprach- und Sprechererkennung haben sich in den vergangenen Jahren wie kaum ein anderer Bereich weiterentwickelt. Erfreulich ist hierbei, dass die Systeme ihren Weg vom Labor in die Anwendung gefunden haben und auch hier unter realen Bedingungen - selbst in störerfüllten Umgebungen - Sprache genügend zuverlässig erkennen. Zahlreiche Anwendungen, die auf Sprachsteuerung basieren, zeugen neben der Entwicklungsreife auch von einer zunehmenden Akzeptanz unter den Benutzern. So ist es nicht verwunderlich, dass viele Anwendungen aus den Bereichen der Sprach- und Sprechererkennung im Ambient Assisted Living Einsatz finden. Im Folgenden findet sich eine Auswahl an Anwendungen für den AAL-Bereich.

Einsatzmöglichkeiten der *Spracherkennung* im AAL-Bereich

- Steuerung von Aktoren (für Fenster / Türen),
- Umfeldsteuerung im Wohnbereich (Licht, Heizung, Jalousien, etc.)
- Abfrage von Videokameras (Innenräume, Haustürbereich),
- Ein- / Ausschalten von Szenarien im Smart-Home
- Akustische Bedienung von Haushaltsgeräten.
- Automatische Auskunftssysteme
- Sprachgesteuerter Rollstuhl
- Blindenhilfe (akustische Bedienung von Geräten)
- Akustische Fernbedienung von TV- und HiFi-Geräten
- Akustische Ziffernwahl beim Telefon
- Akustische Bedienung des Mobiltelefons
- Dialogsteuerung beim Teleteaching/Telelearning
- Akustischer Notschalter bzw. Notrufsystemauslösung
- Diktiersysteme

Einsatzmöglichkeiten der *Sprechererkennung* im AAL-Bereich

- Bankabbuchungen und –überweisungen,
- Haustüröffnung
- Ein- / Ausstellen der Alarmanlage
- Kindersicherung

Einsatzmöglichkeiten der *Sprachausgabe* (Sprachwiedergabe und Sprachsynthese)

- Erinnerungssysteme (Medikamenteneinnahme, Zunahme von Flüssigkeiten)
- Rückkopplungsfunktionen zur Fehlerausgabe, zur Hilfe oder zur Bestätigung von erkannten Kommandos

Akzeptanz der Sprachsteuerung

Grundsätzlich wird die Verwendung von Sprache zur Bedienung bei bewegungseingeschränkten Personen als sinnvoll angesehen. Für diesen Personenkreis stellt Sprache eine Möglichkeit dar, Dinge selbständig zu erledigen, für die andernfalls eine Unterstützung notwendig wäre, oder die nur unter großen Schwierigkeiten bewältigt werden können. Für die Gruppe der älteren Menschen, die zunehmende Einschränkungen im haptischen, visuellen, akustischen und kognitiven Bereich haben ergaben Untersuchungen, das Sprachsteuerungen insbesondere bei Komfortanwendungen als gute Unterstützungsmöglichkeit gesehen werden [2], [6] (vgl. auch Abb.2).

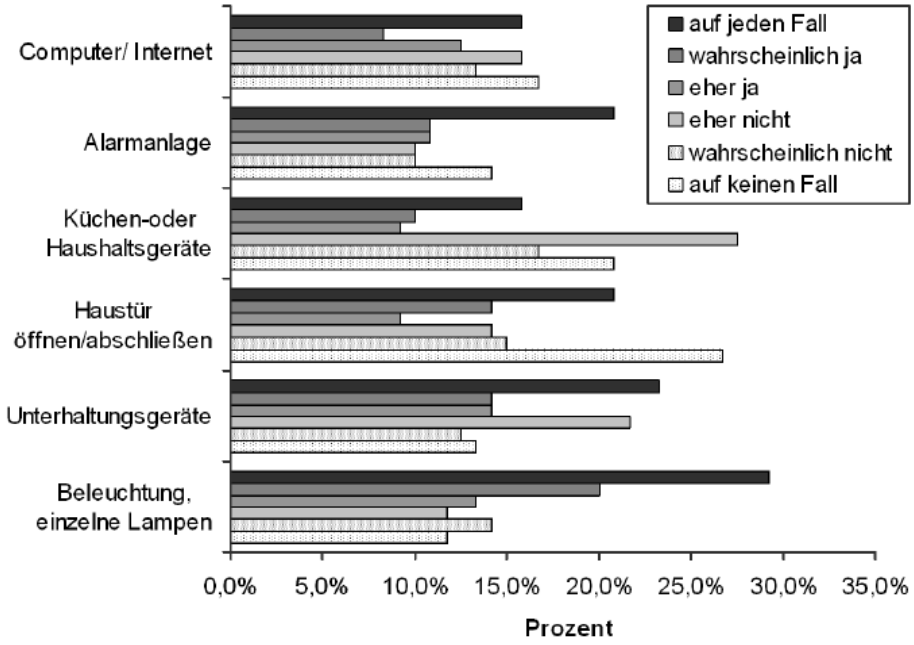

Abbildung 2 – Einsatz von Sprach- und Sprechererkennnung für AAL-Anwendungen für N=120 (Quelle: SENTHA-Survey nach [8])

Zudem wird für ältere Menschen die Sprachsteuerung gerade dann als attraktiv betrachtet, wenn - verglichen mit herkömmlichen Bedienmöglichkeiten - die Benutzung deutlich verbessert wird.

3 Anforderungen für die Nutzung durch ältere Menschen

Die menschliche Stimme wird von Nutzern als selbstverständliches und natürliches Kommunikationsmittel angesehen. Es ist deshalb naheliegend, die natürliche Sprache auch für die Benutzung von Geräten und für Steuerungsaufgaben z.B. im Wohnbereich einzusetzen. Der Einsatz von Sprachsteuerungen sichert damit eine Reihe von Vorteilen, stellt aber

gleichzeitig gewisse Anforderungen an die Systeme, damit eine benötigte Akzeptanz durch die Älteren vorhanden ist. So muss die Sprachsteuerung auch beim freien Sprechen im Raum funktionieren und sollte keinesfalls an ein Headset gebunden sein. Wichtig ist eine zuverlässige Erkennung von Wörtern aus dem alltäglichen Vokabular; eine Liste von vielen Kommandowörtern, die sich der Nutzer alle einprägen muss, stößt auf wenig Akzeptanz. Die Gefahr, ein Kommando zu vergessen und ein Gerät dadurch nicht mehr benutzen zu können, macht es zudem erforderlich, dem Benutzer auch die konventionellen Steuerungsmöglichkeiten offen zu halten. Auch auf lange Trainingsphasen, die zweifelsohne die Zuverlässigkeit verbessern, sollte möglichst verzichtet werden. In der Regel erzielen adaptive Systeme, die sich in der Benutzungsphase an den Sprecher anpassen, ähnlich gute Ergebnisse.

Tabelle 1 – Anforderungen an das System und an den Benutzer

hohe Worterkennungsrate	Das Erkennungssystem sollte eine hohe Erkennungsrate besitzen und eine robuste Erkennung ermöglichen u.a. indem bekannte Störquellen (Dunstanzugshaube, Mixer, Wasserkocher etc.) aus der Anwendungsumgebung mit in die Sprachmodelle einbezogen werden.
Leistungsfähigkeit des Systems transparent machen	Die Erwartungen an die Erkennungsleistungen werden von vielen Nutzern oft als zu hoch eingeschätzt. Es wird vielfach von einer menschlichen Erkennungsleistung ausgegangen, die jedoch aufgrund der fehlenden Informationskanäle (Gestik, Mimik, Tastsinn, Geruch) und der fehlenden menschlichen Erfahrung (z.B. bei Wortdeutungen, fehlenden oder falschen Wörtern) nicht durch die automatische Erkennung erreicht werden kann.
angemessene Trainningsdauer	Das in der Sprach- und Sprechererkennung oft notwendige Training muss angemessen sein und darf nicht zu lang dauern. Für Nutzer ist das Training in der Regel unbequem.
sprecherabhängig oder sprecherunabhängig	Sprecherabhängige Systeme erfordern ein Training, liefern dafür höhere Erkennungsraten. Die Entscheidung, ob sprecherabhängig der –unabhängig hängt von der Anwendung und der Trainingsbereitschaft ab.
Trainingsphase unterstützen	Bei notwendigem Training muss die Trainingsphase sehr einfach vom Benutzer durchführbar sein (z.B. menügeführt mit Sprachanweisungen).
Erkennungsfehler müssen abgesichert sein	Mögliche Fehler bei der Erkennung von Steuerkommandos müssen abgefangen werden, um Falsch-Steuerungen und damit Schäden zu vermeiden.
Sprachdialog; Sprachausgabe als Rückfrage	Sprachdialogsysteme: der Aufbau des Sprachdialogs, insbesondere Nutzung der Rückkopplung, um z.B. einen bestimmten Umfang an Steuerungsmöglichkeiten abzufragen. Kommandowörter können dabei vorgegeben werden und müssen nur nachgesprochen werden.
Kommandowortzahl gering halten;	Anzahl der Kommandowörter sollte 10-15 Wörter nicht überschreiten, sofern der Benutzer sich diese merken muss.
Synonyme verwenden	Neben Kommandowörtern sollten auch Synonyme eingesetzt werden.
System-Aktivierung	Die Aktivierung des Spracherkenners sollte sehr einfach sein und bevorzugt über eine Sprechtaste erfolgen, um die Aktivität des Spracherkenners durch den Nutzer zu kontrollieren (es besteht die Gefahr, das die Sprache außerhalb der Kommandowortgebung eingeschränkt wird, weil der Ältere Angst haben muß, versehentlich ein Gerät einzuschalten).

Für den älteren Menschen ergeben sich folgende Vorteile bei Benutzung der Sprachsteuerung bzw. als Ergänzung zu anderen Benutzungsschnittstellen (vgl. Tab. 2).

Tabelle 2 – Vorteile durch den Einsatz von Sprachsteuerungen

einfaches erlernen, sofortiges nutzen	Die natürliche Sprache ermöglicht ein einfaches Erlernen bzw. den sofortigen Einsatz der Funktionen über die Benutzungsschnittstelle, da eine intuitive Benutzung unterstützt wird.
unabhängig von Handbewegungen und Blick	Hände und Augen sind frei für andere Aufgaben bzw. nicht an die Benutzung gebunden
sofortige Aufmerksamkeit	Für Systemzustände, die einer sofortigen Aufmerksamkeit bedürfen (Alarmsignale, Gefahrenzustände etc.) kann eine verbesserte Wahrnehmung erzielt werden. Akustische Signale können auch wahrgenommen werden, wenn der Benutzer abgelenkt oder unaufmerksam ist bzw. nicht auf das Gerät schaut.
Verhinderung einseitiger Belastung	Es kann eine einseitige Belastung - bei Konzentration der Benutzung auf einzelne Sinnesorgane - verringert werden, indem eine Verteilung und ein Wechsel auf vielfältige Benutzungsmöglichkeiten erfolgt. Der Mensch ist leistungsfähiger, wenn er für gleichzeitige Tätigkeiten unterschiedliche (nicht miteinander konkurrierende) Ressourcen nutzt.
Verbesserung der Benutzung	Die Benutzungsvorgänge können bei eingeschränkter visuell-manueller Interaktion verbessert werden; für hochgradige Einschränkungen z.T. überhaupt erst ermöglicht werden. Die akustische Ein-/Ausgabe kann hierbei eingeschränkte Sinne kompensieren.
Gerätenutzung durch reproduzierbare Laute/Geräusche	Auch bei Menschen mit eingeschränkter Lautbildung (eine Reproduzierbarkeit vorausgesetzt) ist eine Erkennung der Laute bzw. gelernten Geräusche gegeben. In [15] wurde dies genutzt, um einem motorisch Schwerbehinderten eine Notrufauslösung durch ein charakteristisches Geräusch (Zungenschnalzen) zu ermöglichen.
schnelle Eingabe	Bei Diktiersystemen ist eine schnellere Texteingabe gegenüber Tastatur möglich

Insgesamt benötigen sprachbasierte Benutzungsschnittstellen einen vergleichsweise hohen individuellen Anpassungsaufwand – von der Gestaltung der Benutzungsschnittstelle bis hin zu Einbaubedingungen – sind dann jedoch in der Lage, ältere Menschen bei Benutzungsvorgängen zu unterstützen [2], [7].

4 Embedded Systeme

Eine weitere Unterscheidung bei Sprachsteuerungen erfolgt durch die Hardware, wobei in *PC-basierte* und in *integrierte (embedded) Systeme* unterschieden wird.

Einer größeren Bedeutung werden in den nächsten Jahren allerdings den integrierten Systemen zukommen, da die Sprachsteuerung damit nahezu jedem Gerät zugänglich gemacht werden kann.

5 Beispiele für Sprachsteuerungen

Nachfolgend sollen kurz einige embedded Systeme zur Sprachsteuerung beschrieben werden.

Fernbedienung mit integrierter Sprachsteuerung: Untersuchungen im Projekt SENTHA[1] ergaben bei Einsatz der Fernbedienung als Steuergerät für Smart-Home-Funktionen ein

[1] Sentha (Seniorengerechte Technik im häuslichen Alltag) – DFG-Projekt von 1997- 2003

ergonomisches Problem [8]. Einerseits sollten die Smart-Home-Funktionen kontrolliert werden, andererseits aber die Anzahl der Tasten aus Gründen der Akzeptanz weder erhöht noch mit Doppel- oder Mehrfunktionen belegt werden. Eine integrierte Spracherkennung erwies sich hierbei als optimale Lösung, da für die Steuerung lediglich ein Mikrofon und eine zusätzliche Aktivierungstaste für die integrierte Spracherkennung notwendig sind. Damit konnte die Bedienung stark vereinfacht und benutzerfreundlich gestaltet werden. Umfeldsteuerung mit Sicare Pilot (Evosoft/Siemens): Der Spracherkenner *Sicare Pilot* dient insbesondere motorisch behinderten Personen für eine vollständige Steuerung des Smart-Homes über Sprache. Dazu erkennt das System 64 Wörter, die jeweils max. 5 Sprachmuster zum Trainieren der Sprachmodelle benötigen. Die Steuerungsbefehle setzen sich aus mehreren Wörtern z.B. *Haus - Licht1 - an* zusammen und lassen damit eine Steuerung nahezu aller Smart-Home-Komponenten zu. Über eine Sprachwiedergabe werden die erkannten Befehle zur Kontrolle ausgegeben.

Sprachmodul Voice Direct 364 (IBM): Das sprecherabhängige System kann mit bis zu 60 Kommandowörtern trainiert werden. Die Erkennungsrate wird mit bis zu 99% angegeben. Die Chip-Lösung wird durch den Anschluss eines Elektret-Mikrophons und eines Lautsprechers vervollständigt. Die Aktivierung erfolgt durch ein gewähltes Schlüsselwort [9].

vicControl Spracherkennung (VoiceInterconnect): Die Sprachsteuerung vicCONTROL beinhaltet neben einer hochwertigen sprecherunabhängigen Erkennung einschließlich Störgeräuschfilterung auch Komponenten zur Dialogsteuerung und Sprachsynthese für Statusmeldungen. Es können eigene Wortschätze erstellt werden und erfolgreich zur Bedienung von Maschinen und Anlagen eingesetzt [7]. Das System steht für verschiedene Leistungsklassen zur Verfügung (DSP, PDA oder PC).

SpeechAdvance ASR Embedded (SVOX): SpeechAdvance ist ein Spracherkenner für kleine mobile Endgeräte, Automotive- und industrielle Anwendungen. Das System besitzt eine modulare Architektur und kann über eine flexible Konfiguration große Wortschätze und Anwendungen wie sprecherunabhängige Namenwahl oder Zieleingabe in Navigationssysteme abdecken. Das System ist auf Grund seiner geringen Hardwareanforderungen für Smart-Home-Anwendungen, Mobiltelefone, Automotive-Multimedia und PDAs geeignet.

6 Einsatz neuer Technologien zur Systemintegration

Mit dem zunehmenden Einsatz von Sprachsteuerungen in Haushaltgeräten, Wohnumgebungen sowie Geräten zur Unterstützung der Lebensführung werden auch die Ansprüche an die Systeme und die Anforderungen an die Systemintegration der Sprachsteuerungen deutlich größer.

Damit spielen bei den Systemen zunehmend auch technologische Aspekte eine Rolle, um:

- eine höhere Anzahl an Funktionalitäten zu integrieren (Sprach- und Sprechererkennung, Störunterdrückung, Sprachsynthese, Dialogsystem etc.),
- eine höhere Rechenleistung und damit eine schnelle Systemantwort zu ermöglichen,
- eine höhere Miniaturisierung bzw. eine Anpassung an die gegebene Gerätebauformen zu erreichen (flexible, sehr dünne oder auch in Textilien/Kleidung integrierte Systeme),
- energieeffiziente und energieunabhängigere Lösungen bereitzustellen (Energiemanagement und Energieoptimierung und Einsatz von Energy Harvesting Lösungen),
- kostengünstigere Lösungen zu realisieren (Einsatz von Low-Cost-Elektronik, polymerbasierte Systeme, Baukastensysteme) und die
- integrierte Systeme zuverlässiger und damit robuster gegenüber Umgebungsbedingungen (Temperatur, Feuchte, Staub, Vibration) zu machen.

7 Zusammenfassung und Ausblick

Sprachsteuerungen bieten die Möglichkeit ältere Menschen bei der Benutzung von Smart-Home-Anwendungen, technischen Geräten und Assistenzsystemen zu unterstützen. In vielen Fällen kann damit die Benutzung vereinfacht oder sogar erst ermöglicht werden. Ergonomische Aspekte müssen hierbei ebenso Berücksichtigung finden wie anwendungsspezifische Gesichtspunkte, um die nötige Akzeptanz beim Einsatz solcher Lösungen für die Älteren zu erreichen.

Bei den Sprachsteuerungen gibt es noch eine Vielzahl von Lösungsansätzen, um die Interaktion zwischen Mensch und Maschine im Ambient Assisted Living weiter zu verbessern. So werden bereits seit längerer Zeit multimodale Systeme entwickelt und hierbei Gestiken und Mimiken des Menschen mit in die Erkennung einbezogen (vgl. [12], [13]). Zudem werden derzeit Systeme entwickelt, die eine Detektion und Verarbeitung von Wörtern ermöglichen, welche nicht im Vokabular enthalten sind sog. *Out-of-Vocabulary-Wörtern*. Damit ließe sich die fehlerhafte Erkennung weiter reduzieren [16].

Darüber hinaus gibt es derzeit Bemühungen aus der Sprache Informationen über den psychischen Zustand (entspannt, zufrieden, gestresst, panisch etc.) des Nutzers abzuleiten (vgl. auch [10], [11]). In Verbindung mit gemessenen Vitalparametern (u.a. Puls, Hautfeuchte, Temperatur) kann damit auf den Zustand der Person geschlossen werden und schließlich weitere Handlungen daraus abgeleitet werden.

Literatur

[1] In Tagungsband Ambient Assisted Living 1. Kongress, am 30. Jan.-1. Februar, in Berlin, VDI, 2008.

[2] Böhm, U. u.a.: Smart Home für den älteren Nutzer- Akzeptanz, gewünschte Features und Bedienung. Zeitschrift Elektrobörse, (2003). WEKA Fachzeitschriften-Verlag.

[3] van Berlo, A.; Ekberg, J. (eds.): Smart Homes. COST 219bis Publication. Staakes, Helsinki, Finland 1998.

[4] Fellbaum, K.: Ambient Intelligence for a Smart Home Environment. Conference paper 'Intelligent Ambience and Well-Being' Tampere, Finland 19.-20. Sept. 2005.

[5] Hampicke, M.: Optimierung von Smart-Home-Umgebungen für den Wohnbereich. Studientexte zur Sprachkommunikation. Band 31, w.e.b. Verlag Dresden (2005)

[6] Hampicke, M.: Sprachbasierte Benutzungsoberfläche für Smart-Home-Anwendungen. In: K. Fellbaum (Hrsg.): Studientexte zur Sprachkommunikation 20, Cottbus, 2000.

[7] Ploennigs, J.; Kabitzsch, K.; Jokisch, O.; Hirschfeld, D.: Generierung angepasster, sprachbasierter Benutzerinterfaces für die Heim- und Gebäudeautomation. In Tagungsband, 1. AAL-Kongress, am 30. Jan.-1.Februar, in Berlin, 2008.

[8] Mayer, D.; Friesdorf, W.; Heine, A.; [Ed.]. sentha - seniorengerechte Technik im häuslichen Alltag. Springer-Verlag GmbH, 224 S., 2006.

[9] Nauth, P.: Embedded intelligent systems. Oldenbourg Verlag, 2005. XII, 196 Seiten, ISBN 978-3-486-27522-3.

[10] Steidl, S.: Automatic Classification of Emotion-Related User States in Spontaneous Children's Speech. Berlin: Logos Verlag, 2009.

[11] Batliner, A.; Steidl, S.; Hacker, Ch.; Nöth, E.: Private emotions versus social interaction: a data-driven approach towards analysing emotion in speech. In: User Modelling and User-Adapted Interaction - The Journal of Personalization Research, 18 (2008) pp. 175-206.

[12] Engel, R.; Pfleger, N.: Multimodal Fusion. In: W. Wahlster (ed): SmartKom - Foundations of Multimodal Dialogue Systems. 2004.

[13] Porta, D. ; Sonntag, D. ; Neßelrath, R. : A Multimodal Mobile B2B Dialogue Interface on the iPhone. In: MobileHCI '09: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services. September 15, Bonn, Germany, ACM, 2009.

[14] Lawo, M.: Steuerung „Mitdenkender“ Elektroprodukte durch Wearable Computing. In Tagungsband 2. Deutscher AAL-Kongress Technologien – Anwendungen, 27.-28.Januar, Berlin, 2009.

[15] Klaus, H.: Prototyp eines sprachgesteuerten Notrufgerätes für motorisch Schwerbehinderte. In Studientexte zur Sprachkommunikation, Nr.10, ESSV in der Rehabilitationstechnik, Mehnert, D. (Hrsg.), Bln 1993.

[16] BMBF-Projekt *SmartWeb* unter www.smartweb-projekt.de