

# SDV Vol.29.2-2005

SDV. Sprache und Datenverarbeitung  
International Journal for Language Data Processing  
Heft 2/2005

## Inhalt

*Schmitz, Hans-Christian und Christian Weiss*

Introduction (Vol. 2, 2005)

*Scheja, Christian; Ablassmeier, Markus; McGlaun, Gregor and Gerhard Rigoll*

Multimodal, Contextual Dialog Management for Natural Speech in Car Environments

*Hamerich, Stefan W.*

Speech Dialogue Systems for Cars – an Overview

*Vollrath, Mark*

Sprechen und Fahren – Lösung oder Problem?

*Aschenberger, Bianca und Petra Wagner*

A Diagnostic Evaluation of the Speech Input and Output Devices in GPS Navigation Systems

*Goronzy, Silke and Rainer Holve*

A Tool for the Integrated Specification, Design and Implementation of Truly Multimodal Speech-enabled HMIs

*Löffler, Jobst and Martha Larson*

Robust keyword spotting and mobile information retrieval for emergency team communication in the EU Project SHARE

*Oppermann, Reinhard*

Sprache im Kontext multimodaler Information und Kommunikation bei mobiler Computernutzung

## Abstracts

# Introduction (Vol. 2, 2005)

**Autor:**

Schmitz, Hans-Christian und Christian Weiss

**Aufsatztitel:**

Introduction (*Vol. 2, 2005*)

**Jahrgang:**

29

**Heft:**

02 (2005)

**Seiten:**

85-86

Abstract leider nicht vorhanden.

# Multimodal, Contextual Dialog Management for Natural Speech in Car Environments

**Autor:**

Scheja, Christian; Ablasmeier, Markus; McGlaun, Gregor and Gerhard Rigoll

**Aufsatztitel:**

Multimodal, Contextual Dialog Management for Natural Speech in Car Environments

**Jahrgang:**

29

**Heft:**

02 (2005)

**Seiten:**

87-106

**Abstract:**

This paper focuses on an efficient and robust approach to multimodal dialog management which is implemented in a framework for the automotive environment. A dedicated dialog manager for driver's interaction with infotainment and communication systems as well as driver assistance systems has been developed and tested. The dialog manager facilitates

multimodal interaction, i.e. it allows for an efficient integration and combination of natural speech and classical tactile input. The main aim of the dialog manager is to provide flexible and user-centered speech dialogs and to support different interfaces, like buttons or turning knobs. For the dialog control, a frame-based approach is used. The dialog description is realized in XML which allows for a simple overview over the dialog structure. For speech output, a synthetic text-to-speech engine is used. Visual output is displayed on several screens in the car. The usability evaluation shows an improvement of effectiveness, a higher joy of use through the possibility of submitting several pieces of information in only one dialog step with natural speech comparing to a menu-based spoken dialog. The test persons rated this way of frame-based interaction as comfortable and important.

Dieser Beitrag stellt einen effizienten und robusten Ansatz des multimodalen Dialogmanagements im automobilen Umfeld vor. Es wurde ein Dialogmanager entwickelt und evaluiert, der sowohl für Informations- und Kommunikationssysteme als auch für Fahrerassistenzsysteme genutzt werden kann. Dieser ermöglicht multimodale Bedienung auf Intentionsebene. In der aktuellen Version werden natürliche Sprache und taktile Eingaben genutzt. Der Dialogmanager wird genutzt, um ein bestehendes Framework um flexible und nutzer-zentrierte, kontextadaptive, natürlichsprachliche Dialoge zu erweitern. Zur Dialogkontrolle wurde der Frame-basierte Ansatz gewählt. Die Dialogdefinition erfolgt in XML in einem auf die Automobildomäne zugeschnittenen Format. Sprachausgaben wurden von einem TTS-Server synthetisiert. Es wurden daneben mehrere im Auto verbaute Displays angesteuert. In einer Usability-Studie zeigte sich bei den Teilnehmern eine erhöhte Effektivität und Bedienfreude durch die im Frame-basierten Ansatz gegebene Möglichkeit, eine beliebige Anzahl von Informationen in einem Bedienschnitt einzugeben. Dies wurde als sehr komfortabel und wichtig bewertet.

# Speech Dialogue Systems for Cars – an Overview

**Autor:**

Hamerich, Stefan W.

**Aufsatztitel:**

Speech Dialogue Systems for Cars – an Overview

**Jahrgang:**

29

**Heft:**

02 (2005)

**Seiten:**

107-118

**Abstract:**

Speech dialogue systems for cars have been available since 1996. In the last years, speech control evaded more and more into a feature optionally available for nearly every car, not only in luxury vehicles. In this paper, first the roots for speech control made in Germany are discussed. Afterwards, the different types of available speech dialogue systems are presented, followed by explanations of the integration of such systems the technology for in-car speech dialogues is described in more detail. Subsequently, the human-factors of this technology are mentioned and finally two dialogue examples are presented.

Sprachdialogsysteme im Einsatzbereich von Automobilen werden seit 1996 eingesetzt. In den letzten Jahren hat sich Sprachtechnologie mehr und mehr als optionales Zusatzausstattung durchgesetzt und ist nicht mehr nur Autos der Oberklasse vorbehalten. In diesem Beitrag wird zuerst die Sprachsteuerung erläutert, wie sie in Deutschland entwickelt wurde. Anschließend werden unterschiedliche Typen von Sprachdialogsystemen vorgestellt. Dem folgt eine Erläuterung wie solche Systeme eingesetzt und als Sprachdialogsysteme im Auto verfügbar gemacht werden können. Abschließend werden der menschliche Faktor beim Einsatz solcher Systeme betrachtet und zwei Sprachdialogsysteme vorgestellt.

# Sprechen und Fahren – Lösung oder Problem?

**Autor:**

Vollrath, Mark

**Aufsatztitel:**

Sprechen und Fahren – Lösung oder Problem?

**Jahrgang:**

29

**Heft:**

02 (2005)

**Seiten:**

119-134

**Abstract:**

With the rapid development of information and communication technology, the quantity of information for the driver increases. Besides travel related information, driver related information plays an increasing role. In order to process this information and be able to drive safely at the same time, an information management is necessary in any vehicle. Speech interfaces offer a good solution since verbal interactions use resources which are not needed when driving. Thus, the double task “driving and verbal driver-vehicle interaction” should be handy to master. However, in practice, experiments show that verbal interactions may bind attention and use central processing resources which are then no longer available for driving. Thus, for safe driving it is crucially important that verbal interactions are designed in such a

way that for the driver the task „driving“ remains in the focus of attention. In that sense verbal driver-vehicle interactions can solve the problem of increasing information in the vehicle but may also become a problem if this interaction is not optimally designed.

Mit der rasanten Entwicklung der Informations- und Kommunikationstechnologie nimmt auch die Menge an Informationen zu, die für den Fahrer verfügbar ist. Neben Fahrt bezogenen Informationen spielen auch Fahrer bezogene Informationen eine zunehmende Rolle. Um diese Informationsmenge zu verarbeiten und dabei sicher fahren zu können, ist ein Informationsmanagement im Fahrzeug notwendig. Sprachliche Fahrer-Fahrzeug-Interaktionen bieten hier gute Möglichkeiten, da Ressourcen beansprucht werden, die beim Fahren selbst nicht benötigt werden. Die Doppelaufgabe „Fahren und Informationsverarbeitung“ sollte theoretisch gut zu bewältigen sein. Allerdings können auch sprachliche Interaktionen bei entsprechender Auslegung Kapazitäten binden, die für das Fahren nicht mehr verfügbar sind. Entscheidend wichtig für sicheres Fahren ist, dass diese Interaktion so ausgelegt wird, dass für den Fahrer die Aufgabe „Fahren“ auch weiterhin im Vordergrund steht. Insofern ist eine sprachliche Fahrer-Fahrzeug-Interaktion eine Lösung für das Informationsmanagement, kann aber auch zum Problem werden, wenn diese nicht optimal gestaltet wird.

# **A Diagnostic Evaluation of the Speech Input and Output Devices in GPS Navigation Systems**

**Autor:**

Aschenberner, Bianca und Petra Wagner

**Aufsatztitel:**

A Diagnostic Evaluation of the Speech Input and Output Devices in GPS Navigation Systems

**Jahrgang:**

29

**Heft:**

02 (2005)

**Seiten:**

135-146

**Abstract:**

This paper deals with the speech input and output devices of GPS navigation systems, or rather with their quality and user satisfaction. The usage of navigation systems increases steadily, so that the quality assessment and evaluation of the systems itself, but also of their speech in- and output application becomes more and more important. Therefore, we designed and completed an extensive questionnaire and survey, in order to find out the users' opinion, satisfaction and criticism upon these two features. The most important results of this work are presented in this study. Besides, possible weaknesses, suggestions

for improvements and reasonable modifications of the speech input and output applications are derived and discussed.

Der Einsatz von sprachgesteuerten GPS-Navigationssystemen im Automobil steigt ständig an, und mit ihm die Notwendigkeit, die implementierte Sprachaus- und -eingabe qualitativ so hochwertig und robust wie möglich zu gestalten. Dieses Ziel verfolgend, ist zunächst eine umfassende Evaluation und Nutzererhebung hinsichtlich der Sprachapplikationen entscheidend. Sie legt die Bewertungen und Kritikpunkte von Seiten der Endkonsumenten offen, deckt die Schwachpunkte des Systems auf und führt somit zu sinnvollen Verbesserungsvorschläge und zukünftigen Modifikationen. Daher wurde im Rahmen dieser Arbeit eine diagnostische Evaluation durchgeführt, die die Zufriedenheit und Meinungen des Endkonsumenten bezüglich der Sprachaus- und -eingabe von mobilen GPS-Navigationssystemen auswertet. Damit bildet sie die Basis, von den Nutzerbewertungen wiederum auf die Nutzerfreundlichkeit und die verbesserungswürdigen Eigenschaften der Sprachapplikationen zurück zu schließen. Die wichtigsten Ergebnisse sowie die entsprechenden Schlussfolgerungen hinsichtlich der Sprachaus- und -eingabe von GPS-Navigationssystemen werden im Folgenden dargestellt.

# **A Tool for the Integrated Specification, Design and Implementation of Truly Multimodal Speech-enabled HMIs**

**Autor:**

Goronzy, Silke and Rainer Holve

**Aufsatztitel:**

A Tool for the Integrated Specification, Design and Implementation of Truly Multimodal Speech-enabled HMIs

**Jahrgang:**

29

**Heft:**

02 (2005)

**Seiten:**

147-155

**Abstract:**

It is commonly accepted that the use of multiple modalities for human machine interfaces (HMIs) potentially facilitates the handling of complex systems for users. Especially drivers of cars need particularly designed interfaces that do not distract them from their main task – from driving. Therefore, speech dialog systems have been frequently added to the graphical/haptical interfaces because speech control potentially allows drivers to keep their hands on the wheel and their eyes on the street. While speech recognition technology itself was optimised over the past years for the usage in cars so that high recognition rates become

possible even at e. g. high speeds, the overall HMI – i. e. the combination with the graphics/haptics part of the interfaces often lacks consistency and thus ease-of-use. In this paper we present a tool that allows the integrated development of graphical/haptical and speech dialog and that ensures consistent, overall HMIs by using a central, XML-based HMI model.

Es ist allgemein akzeptiert, dass multimodale HMIs dem Benutzer die Bedienung komplexer Applikationen erleichtern. Besonders in einem Fahrzeug werden spezielle Interfaces benötigt, um den Fahrer so wenig wie möglich abzulenken. Aus diesem werden grafisch/haptische Interfaces oft um Sprachdialog erweitert. Während die Spracherkennungsqualität über die Jahre für den Einsatz in Fahrzeugen optimiert wurde, fehlt es Gesamt-HMIs oft an Konsistenz zwischen den Modalitäten. Daher wird die Benutzung für die Benutzer oft eher komplizierter als einfacher. In diesem Artikel stellen wir ein Tool vor, welches die integrierte Entwicklung von multimodalen HMIs ermöglicht. Durch diese integrierte Entwicklung und die Verwendung eines zentralen HMI Modells kann die Konsistenz zwischen den Modalitäten sichergestellt werden.

## **Robust keyword spotting and mobile information retrieval for emergency team communication in the EU Project SHARE**

**Autor:**

Löffler, Jobst and Martha Larson

**Aufsatztitel:**

Robust keyword spotting and mobile information retrieval for emergency team communication in the EU Project SHARE

**Jahrgang:**

29

**Heft:**

02 (2005)

**Seiten:**

157-168

**Abstract:**

The SHARE system is a prototype advanced mobile service that will provide public safety officers with a communication network and on-line decision support during rescue operations and disaster management. The system is being developed within the SHARE project, funded by the EU and led by the Fraunhofer Institute for Media Communication (IMK). The IMK also contributes development effort in the area of speech recognition and retrieval, to which we dedicate the main portion of this paper. A robust keyword spotting system that makes it possible to archive and retrieve spoken communication will be developed and implemented. Speech recognition for public safety communication, especially communication between fire

department officers on which the SHARE project focuses, is a challenging area. The background conditions during emergency incidents include not only standard in-vehicle noise, but also additional noise, such as sirens, horns, traffic and loud voices. Mobile applications require interfaces designed to make access easy and efficient from small end devices. Emergency teams will access the SHARE system using a rugged tablet PC or a PDA on which the information will be presented in an interactive digital map. We show how the system makes it possible to retrieve information from the multimedia database comfortably using such devices.

Das SHARE-System ist Prototyp eines fortschrittlichen mobilen Systems, das öffentlichen Sicherheitskräften ein Kommunikationsnetzwerk und on-line Entscheidungsunterstützung für Rettungseinsätze und Katastrophenmanagement zur Verfügung stellt. Das System wird im Rahmen des Projektes SHARE entwickelt, das von der EU finanziert und vom Fraunhofer Institut für Medienkommunikation (IMK) geleitet wird. Das IMK entwickelt im Rahmen des Projektes Module für Spracherkennung und Information-Retrieval. Der Hauptteil des vorliegenden Artikels ist diesem Thema gewidmet. Ein robustes System zur Schlüsselworterkennung, das in SHARE entwickelt und implementiert wird, ermöglicht die detaillierte Suche im Archiv der indexierten Sprachnachrichten. Spracherkennung für die Audiokommunikation im Bereich öffentlicher Sicherheit stellt eine wesentliche Herausforderung dar. Die Verfahren müssen nicht nur den üblichen Hintergrundbedingungen, die im normalen Fahrzeugbetrieb herrschen, Rechnung tragen, sondern auch mit Geräuschen von zusätzlichen Quellen, wie zum Beispiel Sirenen, Hörnern, Verkehr und laute Stimmen, zurecht kommen. Mobile Anwendung benötigen Schnittstellen, die auch den Zugang mit leicht transportablen Endgeräten einfach und effizient ermöglichen. Sicherheitsteams greifen auf das SHARE System mittels robuster TabletPCs oder PDAs zu, worauf unter anderem auch Informationen in Form von interaktiven digitalen Karten dargestellt werden. Wir zeigen, wie das System bequemen Zugang von solchen mobilen Endgeräten zum Multimedia-Retrieval-Service erlaubt.

# **Sprache im Kontext multimodaler Information und Kommunikation bei mobiler Computernutzung**

**Autor:**

Oppermann, Reinhard

**Aufsatztitel:**

Sprache im Kontext multimodaler Information und Kommunikation bei mobiler Computernutzung

**Jahrgang:**

29

**Heft:**

02 (2005)

**Seiten:**



169-183

**Abstract:**

The paper describes the constraints of mobile use of computers in the process of different contexts of use including stationary and mobile periods. Selection and design of input/output modalities of the user computer interaction have to consider such multiple contexts of use at different places under different usage conditions and together with different primary activities. This contribution discusses the context of use of mobile applications, the multimodality of input and output methods of user computer interaction during mobile usage, and the specific pros and cons of speech during mobile usage. Examples of speech interaction in mobile applications are presented together with chances and design requirements for multimodal user interfaces in several projects.

Der Beitrag beschreibt die Randbedingungen mobiler Computernutzung in einem Ablauf verschiedener Nutzungssituationen, die stationäre und mobile Teilphasen umfassen. In einem solchen Nutzungskontext ist die Auswahl und Gestaltung von Ein-/Ausgabemodalitäten für Inhalte und für Interaktionsmethoden einzuordnen. Entsprechend befasst sich der Beitrag mit dem Kontextbegriff, mit der Multimodalität und mit der speziellen Modalität von gesprochener Sprache im Zusammenhang mit mobiler Nutzung. Es werden Beispiele für Anwendungen mobiler Computersystemen aufgezeigt und die Einsatzmöglichkeiten, Chancen und Gestaltungsanforderungen an multimodale Benutzerschnittstellen anhand konkreter Projekte diskutiert.