

Sprachbedienung im KFZ – Eine Erfolgsgeschichte

Gerhard Hanrieder
Dialog Development
TEMIC Speech Dialog Systems, Ulm, Deutschland
gerhard.hanrieder@temic-sds.com

Abstract: Sprachsteuerung hat in den letzten Jahren auf breiter Front Einzug in die Fahrzeuge gefunden und ist heute in Serie bei vielen Herstellern verfügbar. Das Spektrum reicht dabei von sprachbedienbaren Freisprecheinrichtungen bis hin zu komplexen Infotainmentsystemen inklusive Navigationszieleingabe. Dieser Beitrag beschreibt zunächst am Markt erhältliche Systeme und deren Funktionsumfang. Anschließend werden die benötigten Technologien und Kompetenzen näher beschrieben. Der Beitrag endet mit einem Ausblick auf künftige Entwicklungen.

1 State-of-the-Art

Die Erfolgsgeschichte der Sprachbedienung im KFZ beginnt in der zweiten Hälfte der neunziger Jahre: 1996 führt Mercedes-Benz in der S-Klasse unter dem Produktnamen *Linguatronic* ein Sprachbediensystem für das Autotelefon ein (vgl. [HE01]). Die Erkennung ist sprecherunabhängig und erlaubt Nummern- und Namenwahl sowie einige wenige Kommandos. Innerhalb weniger Jahre wird Sprachbedienung von allen namhaften KFZ-Herstellern angeboten. Gleichzeitig wachsen die Funktionsumfänge der Sprachbedienung (Radio, Kasette, CD-Spieler, Navigation) und die Verfügbarkeit wird von Oberklassefahrzeugen auf die Mittelklasse ausgedehnt. Heute ist Sprachbedienung aus dem KFZ nicht mehr wegzudenken und wird bei der Konzeption neuer Baureihen als integraler Bestandteil des Human Machine Interface (HMI) gesehen. Die Gründe hierfür sind offensichtlich: die stetig wachsende Komplexität der Informations- und Entertainmentprodukte (kurz: Infotainmentsysteme) erfordert ein Bedienkonzept, das es dem Benutzer erlaubt, die Vielfalt der Funktionen auch während der Fahrt sicher und komfortabel zu bedienen. Sprachsteuerung ist hier die erste Wahl.

Die Produktpalette der am Markt verfügbaren Sprachbediensysteme für das KFZ lässt sich in Originalzubehör (OEM = Original Equipment Manufacturer) und Nachrüstmarkt (After Market) untergliedern. Sprachbediensysteme wie das oben genannte *Linguatronic* sind Originalteile, die als – zumeist optionale - Ausstattung beim Händler (hier: Mercedes) geordert werden können. Daneben findet Sprachsteuerung in zunehmendem Maße Einzug in Nachrüstgeräte, wie z.B. Radionavigationsgeräte, aber vor allem in Freisprecheinrichtungen.

1.1 Freisprecheinrichtungen

Sprachbedienbare Freisprecheinrichtungen sind als OEM-Ausstattung oder als Nachrüstlösungen erhältlich. So kann man z.B. die Freisprechlösungen der Temic SDS GmbH aktuell als Option ab Werk bei den Automobilherstellern (z.B. Audi UHV, VW Touareg, etc.) ordern. Darüber hinaus finden sich die Lösungen der Temic SDS in Produkten der Zubehör- und Nachrüstmarktanbieter (z.B. Siemens Car Kit Professional Voice, THB Bury Voice Dial, Funkwerk Dabendorf Audio 3000, Becker Speech Control).

Über die reine Eingabe einer Nummer hinaus stehen bei den sprachbedienten Freisprechanlagen auch Funktionen wie Namenswahl und die Einrichtung sowie Verwaltung eines internen Telefonbuchs zur Verfügung. Dieses Telefonbuch, das unabhängig vom Gerätetyp des verwendeten Handys genutzt werden kann, ist Bestandteil des Systems der Freisprechanlage im Fahrzeug. Mit diesen Funktionen ist es möglich, die wichtigsten Funktionen des Handys während der Fahrt per Sprache zu bedienen, ohne die Hände vom Lenkrad zu nehmen.

Freisprechanlagen enthalten zunehmend eine optionale Bluetooth™-Schnittstelle (Hands-Free-Profile), die zu weiterem Komfort beim Telefonieren im Fahrzeug führt: Durch die kabellose digitale Funkverbindung über Bluetooth™ können das Handy und die Freisprechanlage miteinander verbunden werden, ohne dass hierfür das Mobiltelefon in eine Halterung (Cradle) gesteckt werden muss. Damit entfällt auch die Notwendigkeit einer Handy-spezifischen Halterung.

1.2 Infotainmentsysteme

Moderne Infotainmentsysteme wie etwa das Multi-Media-Interface (MMI) des Audi A8 oder das COMAND-System der E-Klasse von Mercedes-Benz integrieren eine Vielzahl von Informations- und Kommunikationsdiensten (Navigation, Telefon, TV, DVD, etc.).¹ Neben dem haptisch-grafischen Interface verfügen diese Systeme auch über eine Sprachbedienung, die vollständig mit dem haptisch-grafischen HMI verzahnt ist und so eine multimodale Benutzerschnittstelle realisiert.

Die beiden genannten Systeme, die seit 2003 in Serie sind, bieten erstmalig die Sprach-eingabe von Navigationszielen an. Die Eingabe der Orts- und Straßennamen erfolgt dabei mittels Buchstabieren.² Dank eines intelligenten Matching-Verfahrens reicht es aus, nur wenige Anfangsbuchstaben zu sprechen. Eine Liste mit passenden Einträgen wird dann am Display zur Auswahl angeboten. Der in Abbildung 1 gezeigte Dialogauschnitt aus dem Audi A8-System illustriert das Bedienkonzept (U=User, S=System).

Die Eingabe von Navigationszielen per Sprache bedeutet einen wichtigen Meilenstein auf dem Weg zu intuitiv bedienbaren HMIs. Künftige Systemgenerationen werden neben dem Buchstabieren auch die direkte Ganzworteingabe aller Städte- und Straßennamen erlauben und so das User Interface noch natürlicher machen.

¹ Nähere Information hierzu findet sich unter: www.harmaninfotainment.com

² Im COMAND-System können die größten Städte (ca. 800) auch direkt gesprochen werden.

Sprach-Ein-/Ausgabe	Systemreaktion
U: Ziel eingeben	
S: Bitte buchstabieren Sie den Ortsnamen	
U: A U G S	zeigt Liste passender Städtenamen am Display an
S: Bitte wählen Sie den Zielort aus der Liste	
U: die Nummer 1	übernimmt Listeneintrag 1 als Zielort
S: Bitte buchstabieren Sie den Straßennamen	
U: F R I E D B	zeigt Liste passender Straßennamen am Display an
S: Bitte wählen Sie die Straße aus der Liste	
U: die 3 bitte	übernimmt Listeneintrag 3 als Zielstraße

Abbildung 1: Dialogbeispiel Navigationszieleingabe

2 Sprach-Dialog-Systeme

Die erfolgreiche Produktentwicklung von Sprachdialogsystemen und die Integration in unterschiedlichste Systemumgebungen erfordert Kompetenzen und technologische Führerschaft auf zahlreichen Ebenen: Sprachsignalverarbeitung, Spracherkennung, Dialogentwicklung, Systemintegration.

2.1 Sprachsignalverarbeitung und Spracherkennung

Grundlegende Basis eines erfolgreichen Sprachbediensystems ist ein – sprecherunabhängiges – Spracherkennungsprogramm mit hohen Erkennraten. Für die applikations-spezifische Konfiguration des Erkenners sollten komfortable Entwicklungswerkzeuge zur Verfügung stehen, wie z.B. das StarRec™ Grammar Development System [HH00]. Die Verfügbarkeit von Fremdsprachen ist ebenfalls ein wichtiger Faktor.

Um die erforderlichen hohen Erkennraten zu erreichen, ist auch in hohem Maße Know-How der Sprachsignalverarbeitung erforderlich, damit die Erkennung auch bei typischen Störgeräuschen während der Fahrt noch sicher funktioniert. Die Palette der benötigten Verfahren reicht hier von *beam forming* (Mikrofonarrays) über Echokompensation hin zur Geräuschreduktion (Fahrtwind, Regen, Scheibenwischer, etc.).

Natürlichsprachliche Spracherkennung und -verstehen wird derzeit in keinem System im Fahrzeug eingesetzt. Statt dessen wird in einigen Systemen wie z.B. im Audi A8 (siehe auch das Dialogbeispiel in Abbildung 1) Phrasenspotting eingesetzt, um ein gewisses Maß an Natürlichsprachlichkeit zu erreichen [HH00]. Die Praxis hat gezeigt, dass die Dialoge noch nicht komplex genug sind, um den wesentlich höheren Aufwand eines echt natürlichsprachlichen Erkenners zu rechtfertigen. Dies wird sich voraussichtlich erst in der ferneren Zukunft ändern. Im automobilen Umfeld rückt das System nicht ins Zentrum der Aufmerksamkeit des Benutzers. Statt dessen möchte der Benutzer meist mit knappen, kurzen Äußerungen möglichst schnell zum Ziel kommen.

2.2 Dialogentwicklung

Die Entwicklung einer erfolgreichen Sprachdialogapplikation ist sowohl eine technologische als auch eine ergonomische Herausforderung.

Die Berücksichtigung ergonomischer Faktoren in der Designphase ist zentral für die Benutzerakzeptanz. Die Komplexität der Applikationen wird immer mehr zunehmen. Oberstes Ziel des Dialogdesigns muss es sein, eine intuitiv bedienbare Mensch-Maschine-Schnittstelle zu gestalten, die es dem Benutzer erlaubt, das System so effizient und effektiv wie möglich zu bedienen (vgl. [Je02]). Dies bedeutet im Wesentlichen freiere Eingaben (aber nicht notwendigerweise natürlichsprachliche Eingaben) und bessere Hilfestellungen für den Benutzer sowie gute Indikatoren, wann der Benutzer Hilfe benötigt. Sehr wichtig sind in diesem Zusammenhang auch die Evaluierung durch Usability-Tests und Wizard of Oz-Experimente.

Neben einem gelungenen Interface-Design ist das zweite Erfolgskriterium für eine Applikation die fehlerfreie, termingerechte Implementierung: *time to market*. Dies lässt sich nur mit einem klar definierten Entwicklungsprozess erreichen, in dem die Entwickler in den einzelnen Phasen optimal durch Tools unterstützt werden. Bei Temic Speech Dialog Systems wurde zu diesem Zweck das sogenannte *Dialog Development Studio* (DDS) entwickelt. Mit dieser integrierten Entwicklungsumgebung können die Entwickler im Anschluss an die Designphase die Sprachdialoge zügig implementieren und in einer graphischen Simulationsumgebung auf dem PC testen, bevor die Integration auf die Zielsysteme erfolgt. Abbildung 2 illustriert die wichtigsten Komponenten des DDS-Toolkits. Ein zentrales Konzept ist dabei die strikte Trennung applikationsspezifischer Daten von den Verarbeitungsprogrammen, z.B. dem generischen Dialog-Manager (vgl. [Ha98]).

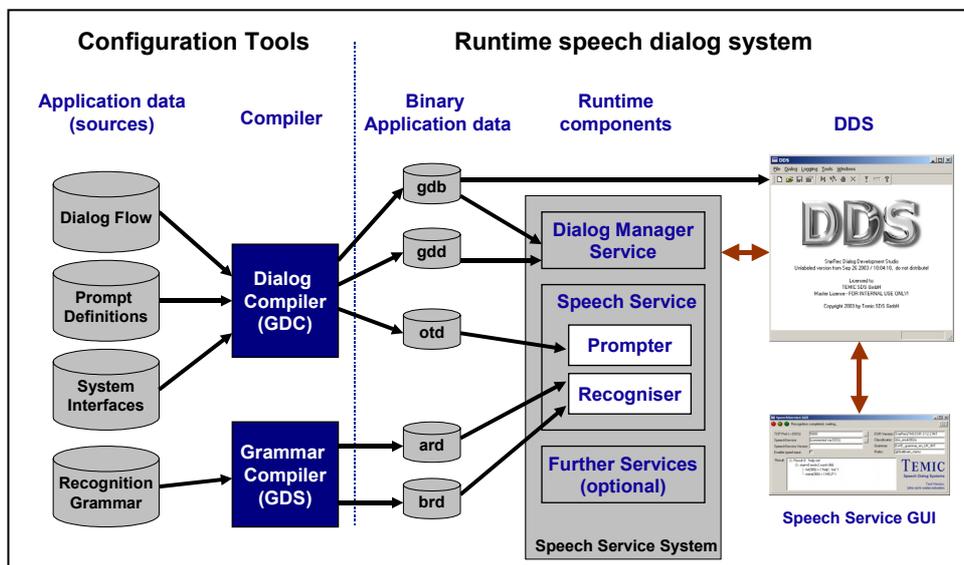


Abbildung 2: Überblick Dialog Development Studio

2.3 Systemintegration

Die Integration von Sprachbediensystemen in KFZ-Umgebungen erfordert breites Wissen und Projekterfahrung im Bereich KFZ-Elektronik. Wenn ein Sprachbediensystem als separate Hardware-Box integriert wird, muss die Kommunikation mit den zu steuernden Geräten über entsprechende Bus-Architekturen erfolgen (D2B, I-BUS, CAN, MOST, etc.). Bei einer reinen Softwareintegration auf einem Infotainmentsystem sind zentrale Faktoren eine schnelle Portierbarkeit auf unterschiedliche Echtzeit-Betriebssysteme (QNX, VX-WORKS, WinCE, etc.) sowie Geräte-Schnittstellen, die leicht an unterschiedliche Kommunikationsprotokolle angepasst werden können. Beim Design der Softwarearchitektur ist deshalb eine klare Trennung generischer Komponenten von plattformabhängigen Schichten zu beachten.

Langjährige Erfahrung in der Produktintegration von Sprachbediensystemen hat uns gelehrt, dass umfangreiches Wissen und Erfahrung im Bereich Systemintegration – mindestens – ebenso wichtig sind wie eine gute Spracherkennung und ein gelungenes Dialogdesign.

3 Fazit und Ausblick

Sprachtechnologie hat sich im KFZ etabliert. Es ist ein Markt vorhanden, der für Systemanbieter lohnend ist. Die Applikationen Telefon und Navigation haben sich dabei als „Killerapplikationen“ der KFZ-Sprachbedienung erwiesen. Als Kernkompetenzen werden Sprachverarbeitungs-, Dialogentwicklungs- und Systemintegrations-Know-How benötigt. Die Systeme werden vom Benutzer akzeptiert, da Fahrsicherheit und Komfort erhöht wird. Die Verbreitung von Sprachbedienung im KFZ wird in den kommenden Jahren noch zunehmen, da sich die Verfügbarkeit von Ober- in Mittelklassefahrzeuge – wie bei anderen Innovationen zuvor – weiter steigern wird und sprachbedienbare Geräte in zunehmendem Maße auch als Nachrüstlösungen erhältlich sein werden. Die Komplexität der Applikationen wird zunehmen, so dass einem intuitiven HMI-Design und der Evaluierung durch Usability-Tests noch stärkere Bedeutung zukommen wird als heute schon.

Literaturverzeichnis

- [HA98] Hanrieder, G.: Integration of a Mixed-Initiative Dialogue Manager into Commercial IVR Platforms. In: Proceedings of IVTTA, Torino, Italy, 1998, pp. 77-82.
- [HE01] Heisterkamp, P.: Linguatronic – Product-Level Speech System for Mercedes-Benz Cars. In: Proceedings HLT, San Diego, USA, 2001.
- [HH00] Hennecke, M.; Hanrieder, G.: Easy Configuration of Natural Language Understanding Systems. In: Workshop on ‘Voice Operated Telecom Services’ (COST 249), Ghent, Belgium, 2000, pp. 51-54.
- [JE02] Jeschke, B.: Design von Benutzerschnittstellen für Sprachbediensysteme im KFZ. In: Tagungsband 13. Konferenz Elektronische Sprachsignalverarbeitung ESSV’02, Dresden 2002.