ELEKTRONISCHE KOMMUNIKATIONSHILFEN MIT SPRACHEIN- UND SPRACHAUSGABE FÜR MOTORISCH BEHINDERTE BENUTZER

Harald Klaus und Klaus Fellbaum

Technische Universität Berlin, Institut für Fernmeldetechnik

Zusammenfassung: *In einem zweijährigen interdisziplinären Forschungsprojekt wurden an der Technischen Universität Berlin elektronische Kommunikationshilfen entwickelt, mit denen motorisch behinderte Benutzer in die Lage versetzt werden, einen herkömmlichen Personal Computer zu bedienen. Schwerpunkt der Arbeiten am Institut für Fernmeldetechnik war hierbei, diesen Personen die Benutzung von handelsüblichen Anwenderprogrammen mit Hilfe von Sprachein- und Sprachausgabesystemen zu erleichtern. In dem vorliegenden Beitrag werden die entwickelten Module vorgestellt und erste Ergebnisse diskutiert.*

1. Einleitung

Eine wesentliche Voraussetzung für die Integration von behinderten Menschen in die Gesellschaft ist ihre berufliche Eingliederung. Durch den zunehmenden Einsatz von Computern in nahezu allen Berufen eröffnen sich auch motorisch behinderten Personen - insbesondere im Büro- und im Kommunikationsbereich - neue berufliche Möglichkeiten. Voraussetzung dafür ist allerdings, daß der Computer über eine behindertengerechte Mensch-Maschine-Schnittstelle verfügt.

Entsprechende Entwicklungen wurden an der Technischen Universität Berlin im Rahmen eines interdisziplinären Forschungsprojektes für motorisch schwer behinderte Benutzer durchgeführt. Die Aufgabe des Instituts für Fernmeldetechnik bestand dabei in der hard- und softwaremäßigen Anpassung von Sprachein- und Sprachausgabemodulen zur Steuerung eines Personal Computers. Mit der Spracherkennung ist der Benutzer jetzt in der Lage, die Kommandos akustisch einzugeben. Die Sprachausgabe wird als Rückkanal verwendet und dient zur Bestätigung, daß das System die Eingabe verstanden hat. Außerdem lassen sich wichtige Informationen auf dem Bildschirm z. B. dadurch wirkungsvoll hervorheben, daß sie vorgelesen werden.

2. Sprachein- und Sprachausgabekomponenten zur PC-Steuerung

Die Sprachsteuerung arbeitet mit dem sprecherabhängigen Einzelworterkennungssystem "VoiceScribe 1000", das von der Firma "Dragon Systems" entwickelt wurde. Aufgrund der hohen Erkennungsrate - sie wird vom Hersteller mit über 99 % angegeben - und einer schnellen Antwortzeit von 0,2 bis 0,3 Sekunden kann das System sehr effizient eingesetzt werden. Für kommerzielle Anwenderprogramme wird in der Regel ein Vokabular von mehreren hundert Wörtern benötigt. Hierfür erwies sich die Möglichkeit des Systems als sehr hilfreich, den Gesamtwortschatz zu unterteilen und die Wortgruppen bei Bedarf zu wechseln. Bei unseren Anwendungen betrug die maximale Größe des Vokabulars weniger als 300 Wörter; die maximale Größe des Vokabulars von 1000 Wörtern ist für die meisten Anwendungen völlig ausreichend.

Die Arbeit am PC beginnt mit einer Login-Prozedur (Bild 1), die zur Identifizierung des Benutzers dient. Ein neuer Benutzer muß zuerst seinen Namen angeben. Anschließend trainiert er das Vokabular mit seiner Stimme. Die Sprachmuster werden auf der Festplatte des PC gespeichert und stehen damit für den späteren Gebrauch zur Verfügung. Der Anwender kann diesen Trainingsvorgang zu jeder Zeit unterbre-

chen und später fortsetzen. Er muß dann lediglich seinen Namen angeben, um die abgespeicherten Sprachmuster in den Arbeitsspeicher zu laden.

Die Login-Prozedur verwaltet die Sprachmuster der einzelnen Benutzer. Dadurch wird gewährleistet, daß das System prinzipiell von beliebig vielen Personen benutzt werden kann. Ihre Anzahl ist lediglich durch die Speicherkapazität der Festplatte beschränkt.

Mit Hilfe des Spracherkennungssystems lassen sich beliebige Anwenderprogramme, beispielsweise Programme zur Textverarbeitung oder zur Verwaltung von Datenbanken, akustisch steuern. Zu jedem Anwenderprogramm muß eine Referenztabelle (Syntax-Overlay) erstellt werden, in der die einzelnen Sprachmuster mit dem jeweiligen Steuerkommando verknüpft werden. Durch einen speziellen Ladezyklus (Bild 2) wird gewährleistet, daß nur diejenigen Programme gestartet werden können, für die bereits Syntax-Overlays vorliegen.

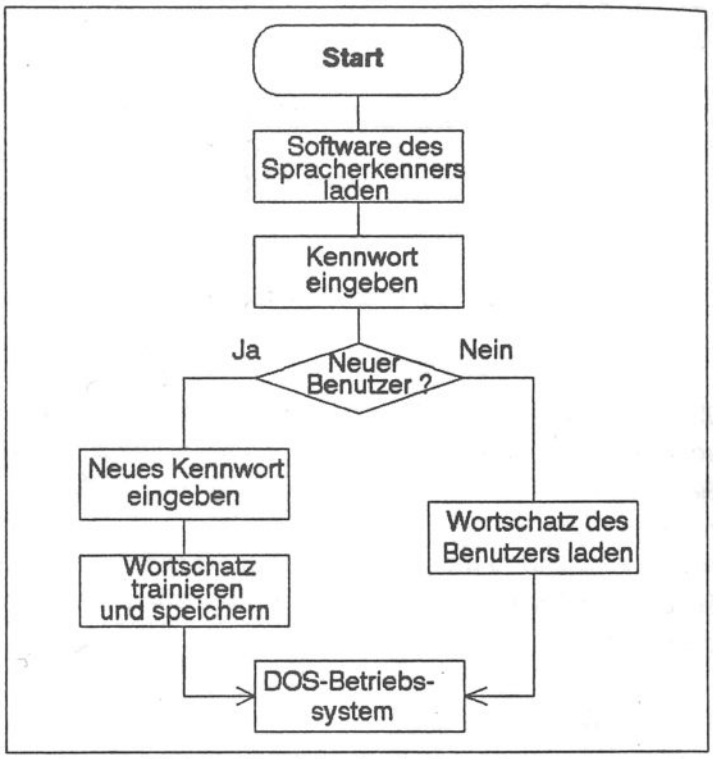

Bild 1: Ablauf der Login-Prozedur.

3. Anwendungen der Sprachkomponenten

Für motorisch behinderte Benutzer ist insbesondere die Handhabung von Büchern sehr mühselig. Zu diesem Zweck wurde eine weitere Benutzerhilfe entwickelt, mit der häufig benötigte Informationen - beispielsweise die Funktionsweise von DOS-Kommandos - auf dem Bildschirm dargestellt werden können. Hierzu gibt der Benutzer einen gewünschten Suchbegriff per Tastatur oder per Spracheingabe ein. Die Information wird einer Textdatei entnommen

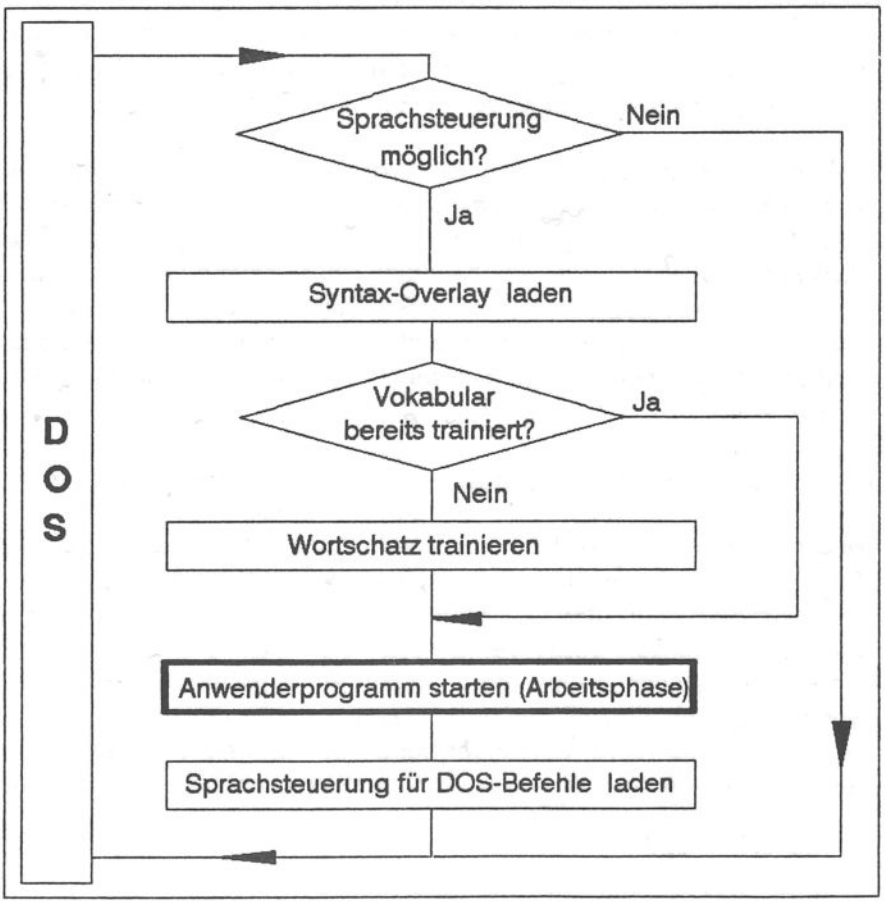

Bild 2: Ladezyklus zum Starten von Anwenderprogrammen.

und seitenweise ausgegeben. Zur Benutzerführung und zur Unterstützung der Bildschirmausgabe wird eine Sprachausgabe verwendet (Bild 3). Dies ist zweckmäßig, da die Sprachausgabe naturgemäß die erwartete Form einer Antwort auf eine gesprochene Eingabe darstellt. Hierdurch kann die Information, die sonst zur Benutzerführung auf dem Bildschirm dargestellt werden muß, oft erheblich reduziert werden. Außerdem lassen sich durch das Vorlesen des Textes wichtige Textpassagen besonders hervorheben.

```
                    Online-Dosinfo-Menü

Sie können zu folgenden Begriffen Hilfeinformationen anfordern:

_ ASSIGN.....Laufwerknamen zuweisen      _ MKDIR.....Katalog erstellen
_ BACKUP.....Festplatte kopieren         _ MODE......Betriebsart festlegen
_ BATCH......Batch-Verarbeitung          _ PATH......Befehlspfad festlegen
_ BREAK......Auf Ctrl-Break prüfen       _ PRINT.....Textdatei prüfen
_ CHDIR......Katalog ändern              _ PROMPT....DOS Eingabeauff. ändern
_ CHKDSK.....Disketten prüfen            _ RECOVER...Fehler beheben
_ COMP.......Dateivergleich              _ RENAME....Dateinamen ändern
_ CONFIG.....Konfiguration definieren    _ RESTORE...Gegenteil von BACKUP
_ COPY.......Dateien kopieren            _ RMDIR.....Katalog löschen
_ CTTY.......Konsole ändern              _ SORT......Daten i.e. Datei sort.
_ DATE.......Datum eingeben/ändern       _ SYS.......DOS übertragen
_ DEL........Dateien löschen             _ TIME......Zeit festlegen/ändern
_ DIR........Katalog prüfen              _ TREE......Kataloge auflisten
_ DISKCOMP...Diskettenvergleich          _ TYPE......Datei anzeigen
_ DISKCOPY...Disketten kopieren          _ VER.......DOS-Version anzeigen
_ FDISK......Festplatte laden            _ VERIFY....Schreiben a. Disk prüfen
_ FIND.......Text in Dateien suchen      _ VOL.......Datenträger anzeigen
_ FORMAT.....Platten initialisieren      _ VERLASSEN.Beenden des Programms

Bitte einen Suchbegriff oder VERLASSEN eingeben:
```

1. Phase: Anzeigen der Überschriften

2. Phase: Sprachausgabe: "Sie können zu folgenden Begriffen Hilfeinformationen anfordern."

3. Phase: Auflisten der DOS-Befehle

4. Phase: Eingabeaufforderung mit Sprachausgabe: "Bitte einen Suchbegriff oder VERLASSEN eingeben!" (mit kurzem Beep).

Bild 3: Ablaufphasen der Benutzerhilfe mit Sprachausgabe.

Weitere Dialogprozeduren wurden entwickelt, um spezielle Hilfe-Informationen anzuzeigen und akustisch auszugeben, ohne daß das Anwenderprogramm verlassen werden muß. Die Information wird in einem Pop-Up-Fenster eingeblendet und - falls gewünscht - durch die Sprachausgabe akustisch ausgegeben. Der Anwender kann die textliche und die akustische Information auf einfache Weise ändern und seinen individuellen Bedürfnissen anpassen.

Bild 4 zeigt das Pop-Up-Fenster der Dialogprozeduren für den Fall, daß der Benutzer im Textprogramm "WORD" Hilfeinformationen zum Laden einer Datei wünscht. Das dreigeteilte Fenster ermöglicht es, bis zu 8 Textzeilen nebst der Überschrift einzublenden. Zusätzlich kann der Benutzer in den alphabetisch sortierten Hilfetextseiten vor- und zurückblättern. Außerdem kann eine Tastaturbelegungstabelle mit den wichtigsten Funktionen sowie ein Indexverzeichnis mit einer Übersicht aller Hilfetexte dargestellt werden.

4. Multimodale Aspekte

Wir konnten erste Erfahrungen mit multimodalen Dialogstrukturen sammeln, die bei der Mischung von Tastatur- und Spracheingaben und bei gleichzeitigen Bildschirm- und Sprachausgaben benötigt werden. Sie zeigten, daß die Optimierung des multimodalen Dialoges von vielen Parametern abhängt und deshalb schwierig ist. So ergab sich z. B., daß eine unerwartete und nicht gewünschte Sprachausgabe eher verwirrend als hilfreich ist, und daß der Umfang der benötigten Hilfeinformation sehr stark von den

```
Tell-me
ÜBERTRAGEN LADEN: Text vom Datenträger holen

Mit diesem Befehl können Sie Texte in den WORD-Editor laden.
Wenn Sie diesen Befehl aktivieren, werden Sie aufgefordert,
den Dateinamen anzugeben. Hierbei wird das Laufwerk und der
Pfad wie in MS-DOS üblich angegeben.
Alternativ zur Texteingabe "zu Fuß" können Sie durch Drücken
der F1-Taste eine Liste der Dateien mit der Erweiterung "TXT"
anzeigen lassen. Wählen Sie nun die gewünschte Datei mit den
Cursortasten und drücken Sie RETURN!

A..Z=Index <Leer>=Übersicht <↵ Hilfetext Bild↑↓=Blättern
+ wiederholtes Sprechen - Sprechen ausschalten <ESC>=Ende
```

Bild 4: Das Pop-Up-Fenster der Online-Benutzerhilfe.

Vorkenntnissen des Benutzers abhängt. Letzteres führt - insbesondere beim Vorlesen des Hilfetextes - bei vielen Benutzern zu einer Über- oder Unterforderung und beeinträchtigt die Akzeptanz des Systems in hohem Maße. Andererseits kann die Sprachausgabe in sehr wirkungsvoller Weise dabei helfen, einzelne Textteile zu kürzen. Sie stellt eine effektive Hilfe dar, wenn sie im Bedarfsfall zugeschaltet werden kann.

5. Schlußbemerkungen

Unserer Meinung nach sind noch viele Aspekte der Mensch-Maschine-Kommunikation ungeklärt. Beispielsweise muß bei Tastatur- oder Spracheingaben näher untersucht werden, in welcher Weise sich der Behinderungsgrad des Benutzers auf die Gestaltung der Mensch-Maschine-Schnittstelle auswirkt, und in welchem Maße die Sprachausgabe die Programmbenutzung beschleunigt und den Bedienkomfort erhöht. Wir hoffen, daß sich einige der Fragen mit den derzeit verfügbaren Systemen (wie dem unsrigen) klären lassen. Was wir hierfür am meisten benötigen, sind intensive Felduntersuchungen mit möglichst vielen (motorisch behinderten) Testpersonen.

6. Literatur

1. K. Fellbaum: "Elektronische Kommunikationshilfen". Weidler-Buchverlag Berlin, 1987.

2. K. Fellbaum, H. Klaus: Abschlußbericht zum IFP "Technische Kommunikationshilfen und Kommunikationstraining zur beruflichen Eingliederung Behinderter. Teilprojekt: Kommunikationshilfen mit Spracheinund Sprachausgabe. TU Berlin, April 1990.

3. M. Fischer: "Erstellung einer menügeführten Sprachsteuerung eines PC (Industriestandard) für motorisch behinderte Benutzer. Diplomarbeit am Institut für Fernmeldetechnik der TU Berlin, 1989.

4. M. M. Taylor, F. Néel, D. G. Bouwhuis: "The structure of Multimodal Dialogue". ELSEVIER Sc. Publ. New York.