

Spracherkennung mit DragonDictate

Luise Widmaier-Müller

DragonDictate ist ein Spracherkennungssystem, das in Verbindung mit einem Computer angewendet wird. Es versetzt diesen in die Lage, gesprochene Sprache zu erkennen und so umzusetzen, daß akustische Befehle befolgt werden und daß das Gesprochene als Text auf dem Bildschirm erscheint. Das System hat selbst nur den Zweck, akustische Signale zu verarbeiten. Es setzt die Existenz eines Textverarbeitungsprogramms voraus, wobei es nicht an eine bestimmte Software gebunden ist. Standardmäßig angeboten wird es für WordPerfect und Microsoft Word 5.5., wofür auch vordefinierte Sprachmakros entwickelt sind.

1. Technische Voraussetzungen

Was man braucht

Das System ist auf einem Rechner installiert, an den ein Mikrophon angeschlossen ist. Es wird für den Einsatz auf MS-DOS-PCs angeboten und kann auf jedem PC mit ausreichender Rechnerkapazität angewendet werden, wobei auch der Einsatz auf einem Mehrplatz-Computersystem denkbar erscheint. Das System benötigt einen Rechner mit einer 80.386 oder 486-Architektur und einen Hauptspeicher mit mindestens 12 Megabyte. Es erfordert eine eigene Festplatte, von der es 20 Megabyte belegt. Das System setzt eine Taktfrequenz von mindestens 33 Megahertz voraus. Eine höhere Frequenz, zum Beispiel von 50 Megahertz, führt zu einer Beschleunigung, die ein schnelleres Sprechen erlaubt und daher mehr Freude an der Arbeit bringt.

2. Start und Training

Für Benutzer mit ein wenig Computererfahrung und Gewöhnung an eine Bedienungsführung ist die Anwendung von DragonDictate relativ einfach.

Eigene Sprachdatei für jeden Benutzer

Am Beginn der Arbeit mit DragonDictate steht das Erstellen einer gesonderten Sprachdatei unter einem eigenen Namen für jeden Benutzer. Hierfür sieht das System ein Trainingsprogramm mit einem Wortschatz von 211 Wörtern vor, der individuell erweitert werden kann um spezifische Wörter, die dem Benutzer von vornherein wichtig sind. Die Wörter werden vom System einzeln vorgegeben und vom Benutzer mehrmals nachgesprochen. Es handelt sich um die Wörter, die praktisch bei jeder Anwendung des Systems benötigt werden, im wesentlichen Menübefehle, Tastatur-, Korrektur- und Formationsbefehle, das Buchstabieralphabet und Interpunktionszeichen. Auf diese Weise lernt das System, Stimme und Aussprache des Benutzers zu erkennen. Dieses Grundtraining erfordert einen Zeitaufwand von weniger als einer Stunde.

Fortgeschrittenes Training

Das Trainieren mittels der gängigen Befehle hat den Effekt, daß diese von Anfang an nahezu hundertprozentig erkannt werden. Das weitere Training geschieht automatisch parallel mit der Erstellung von Texten.

Auf die akustische Vorgabe eines Wortes reagiert das System mit einem Vorschlag. Auf dem Bildschirm erscheint im laufenden Text ein Wort, das das System erkannt hat. Gleichzeitig bietet es in einem Fenster eine Auswahl von klangverwandten Wörtern an. Falls das gesprochene Wort falsch erkannt wurde, kann der Benutzer aus dem Alternativvorschlag wählen, wenn das Wort dort erscheint.

Lehr- und Lernbemühungen ...

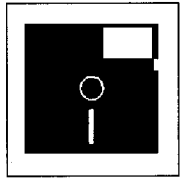
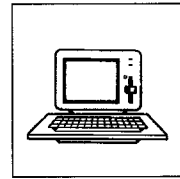
Hat der Benutzer ein bestimmtes Wort korrekt ausgesprochen, wird es jedoch auf dem Bildschirm weder im laufenden Text, noch im Alternativvorschlag angeboten, so korrigiert der Benutzer, indem er nach dem Befehl *neues Wort* das von ihm gewollte Wort buchstabiert. Nach jedem Buchstaben bietet das System neue Wort-Alternativen an, unter denen sich relativ rasch das gesprochene Wort befindet, das dann bestätigt wird. Ab sofort verbindet das System mit dem Klang des gesprochenen Wortes diese Buchstabenfolge, daß heißt, bei identischer Aussprache des gleichen Wortes erscheint dieses in Zukunft sofort auf dem Bildschirm.

... und erste Erfolge

Auf diese Weise lernt das System die Sprache des Benutzers ständig besser kennen. Es paßt die vorhandenen Wortmodelle der Aussprache des Benutzers an, indem es individuelle Wortmodelle erstellt. Die Worterkennungsrates steigt sich so fortwährend.

Um den so mit jeder Arbeit gewonnenen Stand zu erhalten, sieht das System eine Sicherung außerhalb der eigentlichen Textverarbeitung in dem Verzeichnis vor, in dem DragonDictate installiert ist.

Luise Widmaier-Müller ist Rechtsanwältin in Ulm.



Mit dem System kann auf einem Modul eine Vielzahl von Benutzern – jeweils mit eigener Sprachdatei – arbeiten. Die Verwendung des eigenen Benutzernamens gewährleistet das fortlaufende Training und damit die ständige Verbesserung der Worterkennungsrate.

3. Beispiel

Nach etwa fünfstündiger Arbeit der Verfasserin mit dem System spielte sich die Eingabe des Satzes

Der Kläger verlangt von der Beklagten Schadensersatz aus unerlaubter Handlung.
wie folgt ab:

– Die Wörter *Der Kläger* verstand das System sofort.

– Das Wort *verlangt* erkannte es nicht, sondern bot an:

verwandt – verwandelt – verwandten.

Nach Eingabe des Buchstabens *v* bot es an:

verwandt – verwandelt.

Nach Eingabe von *e* wurde *verlangt* angeboten und konnte bestätigt werden.

– Die Wörter *von der* erkannte das System sogleich.

– Für das Wort *Beklagten* bot das System an
beklagte,

nach Eingabe des Buchstabens *B*

Bekleidung – Begabung – Bedeutung,

nach Eingabe des Buchstabens *e*

Befragten – Bekannten – Bedeutung,

nach Eingabe des Buchstabens *k*

Bekannten – Bekanntwerden – Bekämpfung,

nach Eingabe des Buchstabens *l*

Bekleidung – Bekleidungsindustrie,
und weitere fünf verschiedene Ableitungen,

bei Eingabe der restlichen Buchstaben bot das System nichts mehr an, das heißt, das Wort wurde zu Ende buchstabiert und dann bestätigt.

– Die Wörter *Schadensersatz aus* verstand das System.

– Für das Wort *unerlaubt* bot es an
unerwartet – unerwartete,

nach Eingabe des Buchstabens *u,*

unserer – unter – um – ursprünglich – unterstützt,

nach Eingabe des Buchstabens *n,*

untragbar – unmittelbar – und,

nach Eingabe des Buchstabens *e*

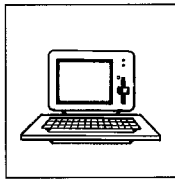
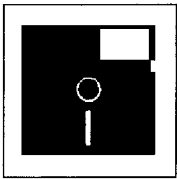
unerlaubte.

Mit dem Befehl, dieses Wort abzuändern, bewegte sich der Cursor an das Wortende. Das System bot dann sofort das gesuchte Wort an nebst den Formen *unerlaubten* und *unerlaubtem*.

– Das Wort *Handlung* wurde sofort erkannt.

Bei einer Wiederholung des vollständigen Satzes erkannte das System jedes Wort sofort.

*Ergebnisse und ihre
Optimierung*



4. Verfügbarer Wortschatz

Nach Herstellerangaben verfügt DragonDictate über ein Lexikon von 80.000 Wörtern. Davon sind bei Installation 25.000 Wörter aktiv und werden vom System erkannt. Weitere 5.000 Wörter können entweder aus dem Lexikon aktiviert oder individuell hinzugefügt werden. Im praktischen Gebrauch läßt sich der individuelle Wortschatz der 5.000 Wörter kaum erschöpfen, da das System das Vokabular ständig aktualisiert, indem es die neuen Wörter aufnimmt und bei Kapazitätserschöpfung die am wenigsten verwendeten Wörter aus dem aktiven Wortschatz löscht.

Ein Test mit juristischen Fachwörtern zeitigte folgendes Ergebnis: Die Wörter *Zurückbehaltungsrecht*, *Kondiktion*, *Kausalzusammenhang*, *Arbeitspflicht* erkannte das System auch bei Eingabe der Buchstabenfolge nicht. Vermutlich sind sie im Thesaurus nicht enthalten. Bei Wiederholungen wurden sie nahezu jedesmal wiedererkannt.

5. Handhabung

Das System setzt nicht nur den zu erstellenden Text von der Sprache auf den Bildschirm um, sondern läßt sich insgesamt akustisch steuern und bedienen. Insbesondere ist auch die Überarbeitung eines Textes, wie Korrektur und Formatierung, akustisch möglich. Bei Erfahrung mit der Handhabung einer Tastatur sind solche Arbeiten allerdings einfacher und schneller über Tastatur auszuführen. Entsprechendes gilt für die Eingabe neuer Wörter, wie oben unter 3. beschrieben, da sich die Buchstaben schneller tippen als mittels des Buchstabenalphabetes eingeben lassen. Die kombinierte Bedienung bietet sich deshalb an.

Die Anwendung von DragonDictate setzt eine gewisse Sprachdisziplin voraus. Das System ist so konzipiert, daß einzelne Wörter erkannt werden, wenn sie durch eine kurze Pause voneinander getrennt werden. Es arbeitet andererseits im Bereich der Sprachbefehle mit der verbundenen Aussprache von mehreren Wörtern. (Beispiel: Der Befehl *neues Wort* ist in einem Wort auszusprechen, da das System andernfalls die Wörter *neues* und *Wort* erkennt und schreibt.)

Das System bietet insgesamt 14 Parameter an, mit denen die Anpassung an die individuellen Benutzungsverhältnisse und den Benutzer verbessert werden kann. Hierzu gehört auch die Anpassung der Länge der Wortpausen an die Sprachgeschwindigkeit des Benutzers.

Da das System über das Mikrophon Geräusche registriert, ist prinzipiell eine ruhige Arbeitsumgebung erforderlich. DragonDictate sieht jedoch einen Parameter zur Anpassung der Mikrophonempfindlichkeit vor. Bis zu einem bestimmten Grad können so ungewollte Signale ausgefiltert werden.

6. Einsatzmöglichkeiten

Einsatzmöglichkeiten bestehen für das System zunächst in allen Fällen und Situationen, in denen ein Text erstellt werden soll, jedoch – aus welchen Gründen auch immer – eine Tastatur nicht benutzt werden kann.

Ein Einsatz erscheint auch dann sinnvoll und effizient, wenn umfangreiches individuelles Schriftgut produziert werden soll, wie es bei der juristischen Tätigkeit insbesondere im forensischen Bereich anfällt. Auch Routineschriftgut, für das nicht auf Textbausteine zurückgegriffen werden kann, läßt sich so rasch erstellen. Hier kann das System zu einer wesentlichen Rationalisierung beitragen.

Weniger effizient erscheint das System bei der Bearbeitung von Texten, die nur geringfügiger Änderungen und Korrekturen bedürfen. Schriftgut, das auf der Basis von Textbausteinen erstellt wird, wie zum Beispiel im Bereich der Vertragsgestaltung, muß individuell und im Detail geändert und ergänzt werden. Es ist relativ aufwendig, durch akustische Befehle den Cursor an die verschiedenen Stellen im Text zu bringen und dort verbal Änderungen vorzunehmen.

7. Dokumentation

Die Dokumentation von DragonDictate ist befriedigend. Es liegt ein deutsches Anwenderhandbuch vor, das die Handhabung und Möglichkeiten des Systems einigermaßen verständlich und konsequent beschreibt. Die Beschreibungen für die akustische Bedienung und für jene mittels Tastatur erfolgen jeweils parallel und sind gestalterisch leicht erkennbar voneinander abgesetzt.

Herstellerangaben ...

... im Praxistest

Akustische Bedienung

Sprachdisziplin

Individuelle Anpassung

Geräuschempfindlich

Wozu das System nützlich ist

Produktion umfangreichen individuellen Schriftguts

Weniger geeignet zur Bearbeitung von Bausteintexten

Befriedigendes Handbuch