

# test

Nr. 1/2004



TEST SPRACHERKENNUNG

COMPUTER + TELEFON

Im Anfang war das Wort. Aber hilft das gesprochene Wort auch beim Umgang mit dem Kollegen Computer? Wir haben Spracherkennungssoftware geprüft und Brauchbares gefunden.

FOTOS: ZEFA / B. RACINE



Reden, reden, nochmals reden! Zuerst heißt es: Die Software trainieren.

## Ein eheähnliches Verhältnis

**A**bschlussbesprechung im Prüfinstitut: Der Leiter fordert die Spracherkennungssoftware mit dem vielleicht schwersten Satz heraus: „Ich spreche jetzt ohne Punkt und Komma – Punkt.“ Das Wort „Punkt“ nach der kurzen Sprechpause ist ein Satzzeichen, die Worte „Punkt und Komma“ soll das Programm als Worte schreiben. In den Wochen der Prüfung war das nie gelungen. Die Spracherkennungsprogramme warfen stur Satzzeichen aus. Doch nun sahen wir – Premiere – diesen Satz völlig korrekt auf dem Monitor. Das Programm hatte gelernt, es war ein Stück weiter gekommen. Oder aber der Sprecher. Der Mensch ist weit flexibler als die Technik. Er passt seine Sprechweise den Macken der Spracherkennungssoftware an. Er spricht sauberer und mit deutlichen Pau-

sen vor Steuerbefehlen für Satzzeichen, Zeilenumbrüche und dergleichen. Ähnlich wie bei einer guten Partnerschaft lernen auch hier beide Seiten, sich aufeinander einzustellen.

### Die zwei Siegertypen

Das Fazit unserer Prüfungingenieure: Nach der unvermeidlichen Übungsphase, durchaus als „Tal der Tränen“ zu bezeichnen, sind die Programme linguattec Voice Pro 10 USB-Edition (beste Erkennungsrate) und IBM Via Voice 10 (nicht ganz so lernfähig wie linguattec) gut brauchbar. Die anderen Programme halten in der Leistung und manchmal auch in der Ausstattung nicht ganz mit den genannten Siegertypen mit. Über alle Prüfpunkte hinweg ist vor allem der Rückstand von VoiceOffice deutlich. Obwohl im Pro-

grammkern eng mit IBMs Via Voice verwandt, ist es keine gute Hilfe. Weniger wegen der Erkennungsleistung, die auch nicht überzeugt. Vor allem jedoch wegen seiner Bedienungs Macken. Mal funktioniert der Hilfe-Button nicht (Anklicken bringt nichts), mal ist ein Korrekturfenster (zum Lernen eines nicht erkannten Wortes) viel zu klein, um den Begriff einzutippen. Die Tabelle auf der Seite 40 gibt einen Überblick.

Vier Programme haben eine „gute“ Spracherkennung. Sie helfen allen:

- die diktieren und beide Hände frei haben müssen – Mediziner beispielsweise;
- die viel mit Standardtexten arbeiten – wie Rechtsanwälte und Steuerberater;
- die behindert sind und mit Maus und Tastatur nicht gut umgehen können.
- die schreibfaul sind.



Wie in einer guten Partnerschaft stellen sich beide Seiten, hier Programm und Benutzer, aufeinander ein.



#### KURIOSSES AM ANFANG

### „Kulturrummel Nation“

Die geprüften Programme funktionieren nach einer Einübungsphase recht gut. Dennoch verstanden und schrieben sie auch manch Kurioses:

„**Kulturrevolution**“: Kulturhunderts flohen, Kulturrummel Nation, Kultur Null Nation, kulturelle Nutzung.

„**Recht auf Ruhe**“: Recht auf Ohren, rechte Ruhe, recht Aufruhr.

„**Zwecks Terminvereinbarung**“: Sex-terminvereinbarung.

„**Ausgelaugt**“: ausgeraubt.

„**Übermüdet**“: über Mütter der.

„**Tippen**“: retten, töten, finden, geben.

„**Euro**“: Otto, Eure, Ego.

„**Johann Wolfgang von Goethe**“: Juan Wolfgang von Göbel, Juan-Wolf kam Mohnblüte, Johann-Wolfgang Wohngüte.

„**Formatier das fett**“ (ein Befehl): von Mathias Fett, Komma („“) Tier das Fett, formatier das Feld.

*Die Erkennungsrate steigt, wenn falsch umgesetzte Wörter und Befehle sofort wiederholt, gegebenenfalls buchstabiert werden. Es hilft auch, dem Programm typische Textdokumente einzugeben. Es stellt sich bei der Analyse dieser Texte auf die Stilebene und den Wortschatz des Anwenders ein.*

Obwohl die geprüften sechs Programme auf zwei Grundmodulen basieren (Dragon hat ein eigenes Spracherkennungsmodul, alle andere nutzen Versionen von IBMs ViaVoice), richten sie sich doch an verschiedene Zielgruppen: IBMs Via-Voice und linguattec Voice Pro bieten die beste Spracherkennung. Für den professionellen Einsatz ist noch wichtig: Kann ein Fachwortschatz geladen und können Audiodateien vom Diktiergerät eingespeist werden? Wie lernfähig ist das Programm? Und wie resistent ist es gegenüber Hintergrundgeräuschen?

#### Erkennen, navigieren, lernen

In der Summe der Eigenschaften reduziert sich die Auswahl auf IBM ViaVoice Pro 10 und linguattec Voice Pro 10. Das Paket linguattec gibt es derzeit als spezielles

Angebot für Mediziner mit einer Fachwortsammlung zu zehn medizinischen Fachgebieten. Es kostet knapp 400 Euro.

Mancher ist jedoch mehr auf die Programmsteuerung (Navigieren) durch Spracheingabe angewiesen und kann mit etwas schlechterer Spracherkennung beim Diktieren leben. Wir prüften das mit Word, mit dem Internet Explorer und dem Mail-Programm „Pegasus“. Die Dragon-Programme schnitten am besten ab.

Doch vor dem Erfolg muss viel Fleiß und Zeit investiert werden. Zuerst ist ein vorgegebener Text zu sprechen, damit das Spracherkennungsprogramm die ihm bekannten Wörter mit unserer Aussprache verbinden kann. Das dauert bis zu 15 Minuten. Wenig hilfreich war das zusätzlich von den Programmen angebotene unspezifische Training. Dazu wird

ein weiterer Text gesprochen. Der war zwar lustig (bei IBM ein kritischer Exkurs über Computer und ihre Macken), doch die Erkennungsrate stieg nicht: Eine halbe Stunde und jede Menge Wasser zum „Ölen“ der trockenen Kehle waren nutzlos vertan. Weit hilfreicher fanden wir die anderen Lernoptionen:

- Der Korrekturmodus, bei dem ein nicht erkanntes Wort eingetippt und gegebenenfalls noch einmal gesprochen wird. Das halbierte die Fehlerrate etwa und dauerte selbst zu Anfang bei einem längeren Geschäftsbrief kaum mehr als eine Viertelstunde zusätzlich. Danach sank der Aufwand spürbar.
- Der Buchstabiermodus, bei dem ein falsch erkanntes Wort Buchstabe für Buchstabe vorgesprochen und später stets korrekt erkannt wird.



**linguattec Voice Pro 10 USB-Edition**  
199 Euro

**GUT (1,9)**

Beste Spracherkennung und Lernfähigkeit, besonders resistent gegen Hintergrundgeräusche. Vielseitig genug auch für professionellen Einsatz, unter anderem mit Audiodatenimport (Diktiergerät) sowie umfangreichem, erweiterbarem Fachwortschatz.



**IBM ViaVoice R10 Pro Edition**  
90 Euro

**GUT (2,0)**

Gute Spracherkennung und besonders resistent gegen Hintergrundgeräusche. Vielseitig genug auch für professionellen Einsatz, unter anderem mit Audiodatenimport (Diktiergerät) sowie erweiterbarem Fachwortschatz.



**Dragon Naturally Speaking 7 Standard**  
99 Euro

**GUT (2,3)**

Sehr ähnlich dem Dragon Preferred, jedoch nicht so vielseitig: unter anderem kein Audiodatenimport (Diktiergerät), Fachwortschatz nicht erweiterbar. Noch „gute“ Spracherkennung und „gut“ nutzbar für das Navigieren (Programmsteuerung durch Sprache).



**Dragon Naturally Speaking 7 Preferred**  
179 Euro

**GUT (2,4)**

Sehr ähnlich, jedoch besser ausgestattet (vielseitiger) als das Dragon Standard. Schlechteste Erkennungsrate im Test, jedoch „gut“ nutzbar für das Navigieren (Programmsteuerung durch Spracheingabe).



**IBM ViaVoice Standard Edition R 8.02**  
50 Euro

**GUT (2,4)**

Noch „gute“ Spracherkennung und besonders resistent gegen Hintergrundgeräusche. Jedoch recht geringer Befehlsumfang für das Navigieren (Programmsteuerung durch Spracheingabe).



**Voice Office Premium 10**  
179 Euro

**BEFRIEDIG. (3,1)**

Erkennungsrate nur nach einem zusätzlichen Lernprozess akzeptabel, schlechtestes Programm beim Navigieren (Programmsteuerung durch gesprochene Befehle).

### Spracherkennungssoftware

	linguattec Voice Pro 10 USB-Edition	IBM ViaVoice R10 Pro Edition	Dragon NaturallySpeaking Version 7 Standard	Dragon NaturallySpeaking Version 7 Preferred	IBM ViaVoice Standard Edition Release 8.02	Voice Office Premium 10
Mittlerer Preis in Euro ca.	199	90	99	179	50	179
<b>test-QUALITÄTSURTEIL</b>	<b>GUT (1,9)</b>	<b>GUT (2,0)</b>	<b>GUT (2,3)</b>	<b>GUT (2,4)</b>	<b>GUT (2,4)</b>	<b>BEFRIEDIGEND (3,1)</b>
<b>SPRACHERKENNUNG</b> 50%	<b>gut (2,0)</b>	<b>gut (2,1)</b>	<b>gut (2,4)</b>	<b>befriedigend (2,9)</b>	<b>gut (2,5)</b>	<b>befriedigend (2,7)</b>
Aufsatz und Geschäftsbrief	+	+	○	○	+	○
Gedicht	○	○	+	○	○	○
Einfluss von Hintergrundgeräuschen	++	++	+	+	++	○
Lernfähigkeit	++	+	+	⊖	○	++
Navigation	○	○	+	+	⊖	-
Technische Prüfungen	+	+	⊖	⊖	-	+
<b>HANDHABUNG</b> 40%	<b>gut (2,1)</b>	<b>gut (2,0)</b>	<b>gut (1,9)</b>	<b>gut (1,9)</b>	<b>gut (2,0)</b>	<b>ausreichend (3,6)</b>
Benutzerführung	+	+	+	+	+	⊖
Handbuch und Hilfe	+	+	+	+	+	⊖
Installation und Deinstallation	+	+	+	+	+	○
<b>VIELSEITIGKEIT</b> 10%	<b>sehr gut (1,0)</b>	<b>sehr gut (1,5)</b>	<b>ausreichend (3,6)</b>	<b>gut (2,4)</b>	<b>befriedigend (3,0)</b>	<b>gut (2,5)</b>
<b>TECHNISCHE MERKMALE/AUSSTATTUNG</b>						
Audiodatenimport (Diktiergerät)	■	■	□	■	□	■
Nachträgliches Abhören des Diktats	■	■	□	■	■	■
Speichern des Diktats als Audiodatei	■	■	□	□	■	■
Fachvokabular mitgeliefert	Recht, Wirtschaft, Computer, IT	□ <sup>1)</sup>	□	□	□	Wirtschaft
Erweiterungen für Fachvokabulare erhältlich	■	■	□	□	□	■
Sprachgesteuerte Maus-/Tastaturbedienung	■/■	■/■	■/■	■/■	□/□	□/■
Vorlesen mit synthetischer Stimme	■	■	□	■	■	■
Headset im Lieferumfang/USB Headset	■/■	■/□	■/□	■/□	■/□	■/■
Besonderheiten			Wortschatz einsehbar, kein Im-/Export von Benutzerdaten	Wortschatz einsehbar, Integriertes englisches Vokabular		Knappes Handbuch; kein eigenes Texteingabefeld (Pad)
<b>Bewertungsschlüssel der Prüfergebnisse:</b> ++ = Sehr gut (0,5-1,5), + = Gut (1,6-2,5), ○ = Befriedigend (2,6-3,5), ⊖ = Ausreichend (3,6-4,5), - = Mangelhaft (4,6-5,5).						
<b>Bei gleicher Note Reihenfolge nach Alphabet. Prozentangaben = Gewichtsanteil am test-Qualitätsurteil.</b>						
■ = Ja. □ = Nein. Anbieter siehe Seite 100. <sup>1)</sup> Registrierte Kunden erhalten kostenlos CD mit Fachwortschatz Recht, Wirtschaft, Computer, IT und allgemeine Medizin.						



## ZUBEHÖR

## Damit können Sie auch sprechen

**Sprechgarnituren:** Zu den Spracherkennungsprogrammen gehören Sprechgarnituren, „Headsets“, die im Test allesamt ordentlich funktionierten. Wir haben auch ein teureres nachgekauft, das aber kaum Vorteile gegenüber den mitgelieferten zeigte. Wer dennoch ein neues Headset sucht, findet im Internet unter <http://support.scansoft.com> für Dragon Naturally Speaking und IBM ViaVoice compatible Modelle.

**Diktiergeräte:** Bis auf die (billigen) Standardversionen von Dragon und IBM können alle geprüften Spracherkennungsprogramme Audiodaten importieren. Mit etwas Ausprobieren erkennen sie die vom Diktiergerät (oder von Taschen-PCs wie dem iPaq von Compaq) importierten Sprachnotizen genauso wie das, was der Benutzer in die Sprechgarnitur diktiert hat.

**Im Test:** 6 Spracherkennungsprogramme. Einkauf der Prüfmuster: September 2003. Preise laut Anbieterbefragung im Oktober 2003. Die Prüfpunkte Spracherkennung, Installation und Deinstallation sowie Benutzerführung wurden auf insgesamt vier vergleichbaren IBM-kompatiblen Rechnern mit 256 MB RAM, Taktraten von 1,2 bis 2,2 GHz und dem Betriebssystem Microsoft Windows XP Home Edition getestet.

### Spracherkennung ► 50 %

Zum Testen der Spracherkennung diktierten zehn Prüfpersonen, davon zwei mit Dialekt, einer mit Akzent und ein Kind, bei jedem Prüfmuster drei **Aufsatztexte**, einen **Geschäftsbrief** und ein **Gedicht**.

Außerdem wurden mit vier Testpersonen weitere Tests zum **Einfluss von Hintergrundgeräuschen** bei einem mittleren Schalldruckpegel der Störgeräusche am Ort des Einsprechmikrofons von 56 dB(A) durchgeführt.

Die **Lernfähigkeit** der Software wurde von vier

### UNSER RAT

Für die **Spracherkennung im Heimgebrauch** genügt das gerade noch „gute“ **IBM ViaVoice Standard R. 8.02**. Es kostet 50 Euro – weniger als alle anderen Programme im Test.

Lernfähiger und vielseitiger sowie mit gesondertem, größtenteils sogar schon mitgeliefertem Fachvokabular aufzuwerten und deshalb richtig für den **professionellen Einsatz** ist **linguatec Voice Pro** für 199 Euro.

Wer das Spracherkennungsprogramm vorwiegend zur **Bedienung des Computers** (Navigieren) statt für das Diktieren von Texten nutzen will, bekommt mit **Dragon NaturallySpeaking 7 Standard** (99 Euro) eine gute Unterstützung.

Die Dokumentenanalyse („an den Schreibstil anpassen“ oder „Vokabularanalyse“ genannt). Dabei durchforstet das Programm ein oder mehrere Dokumente. Es wirft Wörter aus, die sein Wortschatz nicht enthält (die werden ihm dann vorgesprochen) und stellt sich auf die vom Benutzer häufig verwendeten Wortgruppen und seinen Satzbau ein.

### Vor Stilblüten nicht gefeit

Trotz aller Lernerfolge bei der Software und ihrem Besitzer – ein völlig fehlerfreies Erkennen des gesprochenen Wortes und 100 Prozent korrektes Schreiben



FOTO: B. RACINE

Du verstehst mich!

sind nicht zu erwarten. Zu Erkennungsfehlern gesellen sich überraschende Rechtschreib- und viele Grammatikfehler. Tückisch sind vermeintlich richtige, tatsächlich jedoch falsch erkannte Wörter. Das Programm macht darauf nicht aufmerksam. Es schreibt einfach vor sich hin. Aus „Kulturrevolution“ wurde beispielsweise „Kultur Null Nation“. Je lyrischer der Text, desto mehr Stilblüten gab es (ganz schlimm beim Vortrag des Gedichts „Der Erlkönig“). Und als wir vorlesen, dass „Mediziner nun ein Recht auf Ruhe haben“, postulierte das Programm ein „Recht auf Ohren“.

### AUSGEWÄHLT GEPRÜFT BEWERTET

Testpersonen durch Korrigieren der aufgetretenen Fehler und wiederholtes Diktieren ermittelt. Bei der **Navigation** testeten drei Prüfpersonen typische, durch Spracheingabe gesteuerte Vorgänge, die bei der Nutzung des Dokumentenverarbeitungsprogramms Microsoft Word 2000, dem Browser Microsoft Internet Explorer 6 und dem E-Mail-Programm Pegasus Mail 4 auftreten.

Bei den **technischen Prüfungen** untersuchten wir die Start- und Beendungsdauer sowie den Ressourcenbedarf auf der Festplatte und im Hauptspeicher. Ferner beurteilten wir die Bearbeitungszeit, den Speicherbedarf und die CPU-Zeit für eine vorgegebene Aufgabenstellung.

### Handhabung ► 40 %

Zur Beurteilung der **Benutzerführung** urteilten ein Fachmann und zwei interessierte Laien über die Handhabbarkeit der Software, die Einheitlichkeit der Begriffe und des Layouts, die Verständlichkeit von Fehlermeldungen und die Übersichtlichkeit der Bedienoberfläche. Dabei wurden die Prüfpunkte

Programmstart, Menü-Navigation, Hilfsfunktionen, Fehler- und Warnhinweise und Korrekturphase untersucht.

Ein Fachmann und ein interessierter Laie beurteilten die **Handbücher und Hilfsfunktionen** der Softwarepakete. Dazu wurden auch die Angaben der Systemvoraussetzungen und des Funktionsumfangs überprüft. Ferner wurden das äußere Erscheinungsbild, der Aufbau und die Gliederung der Handbücher sowie die Handhabung der Handbücher und der Hilfsfunktionen geprüft. Abschließend wurde die Übereinstimmung zwischen Dokumentation und Software kontrolliert.

Bei der **Installation und Deinstallation** wurden die Angaben in der Dokumentation, die Beschreibung der Datenträger sowie die Durchführung der Installation und der Deinstallation untersucht.

### Vielseitigkeit ► 10 %

Bei der Vielseitigkeit bewerteten wir insgesamt 23 Ausstattungsmerkmale.