

Der feine Luxus des schnellen Diktats steht im Zeichen des Drachen

Die neue Spracherkennung Dragon Naturally Speaking 11 setzt Maßstäbe hinsichtlich der Erkennungsgenauigkeit. Diktieren ist immer schneller als Tippen.

Von Michael Spehr

Nach ungefähr zwei Stunden stellen sich die ersten Aha-Effekte ein, und dann kommt man aus dem Staunen nicht mehr heraus. Dass man sich nicht nur oberflächlich mit dieser Software beschäftigen muss und eine gewisse Übung benötigt – das ist der größte Nachteil der Spracherkennung am Computer. Wir haben es in den vergangenen Jahren immer wieder erlebt: Wer mit der Technik allein herumexperimentiert, ungeduldig ist und elementare Regeln missachtet, gibt nach kurzer Zeit auf. Wer jedoch Spracherkennung bei einem erfahrenen Benutzer im praktischen Einsatz sieht, will sie haben, die Faszination ist groß.

Spracherkennung wird von Jahr zu Jahr besser, derzeit hat sie ihren Durchbruch auf den modernen Smartphones. Mit dem offenen Betriebssystem Android in der aktuellen Variante 2.2 und mit dem iPhone von Apple kann man SMS und andere Nachrichten diktieren, die passenden „Apps“ vorausgesetzt. Die beste und zuverlässigste Lösung mit hoher Erkennungsra-

te ist „Dragon Dictation“ für das iPhone. Das Mini-Programm steht gratis parat, der Hersteller Nuance will damit für seine PC-Spracherkennung Dragon Naturally Speaking werben. Dass dies gelingt, zeigt die Platzierung der App als eine der meistgeladenen in Apples Online-Geschäft.

Nun bringt Nuance seine Windows-Software in der Version 11 auf den Markt und zeigt, was sich in den vergangenen zwei Jahren getan hat: Das vielfach verbesserte Programm ist der Maßstab, es gibt keine andere Software mit vergleichbarer Leistungsfähigkeit, man mag diese Monopolstellung bedauern. Wir haben den neuen Drachen zunächst an seiner Erkennungsleistung gemessen: Wie viel Prozent der Worte werden richtig erkannt? Das hört sich trivial an, ist es aber nicht. Berücksichtigt man, dass eine Spracherkennung nur jene Begriffe fehlerfrei umsetzen kann, die in ihrem Vokabular gespeichert sind, hängt die Beurteilung von den Texten und Inhalten ab. Eine prägnante Zusammenfassung der Art „98 Prozent Genauigkeit“ ist also unsinnig. Ein Arzt, Anwalt oder Gutachter, der für seine Diktate ein begrenztes Fachvokabular verwendet und dies sinnvollerweise gleich mit Dragon dazukauf, erreicht spielend 99 Prozent und mehr. Ein Nutzer mit einem reichhaltigen Vokabular, sagen wir ein Schriftsteller oder Journalist, wird eher bei 97 Prozent liegen. Er wird vermutlich eigene Ad-hoc-Begriffe (wie diesen hier) verwenden und vielleicht nur einmal und dann nie wieder. Kein Wörterbuch kann da mithalten.

Wie gut die Erkennungsleistung von Dragon 11 im Vergleich mit der Vorgängerversion 10 ist, haben wir mit unseren eigenen Diktaten ermittelt, das ist eine subjektive Komponente im Test. Aber wir suchten reproduzierbare Ergebnisse. Deshalb wurde aus Hunderten von Diktiergerätaufzeichnungen im DSS-Pro-Format eine Auswahl erstellt. Neben kurzen, knappen E-Mails auch längere mit privatem Inhalt, Nachrichtenmeldungen, ein Protokoll einer Testfahrt im Auto und spaßeshalber ein wissenschaftlicher Aufsatz aus der DDR-Geschichtsschreibung in einer aus heutiger Sicht merkwürdigen Sprache. Ein und dieselbe Diktat-Datei ließen wir der Reihe nach erkennen, und zwar mit der alten und der neuen Version, so wie sie der Nutzer nach der Installation vorfindet (also ohne benutzerspezifische Anpassungen). Die Ergebnisse: Die Fehlerrate ist in etwa halbiert, die Genauigkeit liegt – je nach Text – bei bis zu 99 Prozent, ein vorzügliches Ergebnis. Es fußt auf zwei Verbesserungen: Zum einen ist das Vokabular mehr als doppelt so groß. Rund 300 000 Einträge zählt das „Bestmatch IV“ der aktuellen Version, 135 000 waren es bei „Bestmatch III“ in Dragon 10. Zum anderen wurde die Abtastrate des Mikrofons von 11 auf 22 Kilohertz ebenfalls verdoppelt, und so verschwinden viele Probleme mit undeutlich gesprochenen Wortendungen quasi von allein.

Schon wegen des größeren Vokabulars empfiehlt es sich nicht, ein altes Dragon-Profil für die neue Version zu aktualisie-

ren. Man fange von vorn an. Bei uns arbeitete das neue 11er-Profil besser als das zwei Jahre lang gepflegte und persönlich angepasste 10er-Profil. Gegebenenfalls exportiere man seine Wortliste und importiere sie in die neue Version. Das gilt aber nur für Anwender, die ein eigenes opulentes Fachvokabular aufgebaut haben. Die typischen Schwierigkeiten sind übrigens in beiden Versionen gleich: Groß- und



Kleinschreibung, Getrennt- und Zusammenschreibung, Abkürzungen, Produkt- und Eigennamen bereiten noch immer Kopfzerbrechen. Mit ein bisschen Erfahrung kennt man im Laufe der Zeit die Stolpersteine und wird dann beispielsweise „ist vom Feinsten“ dahingehend diktieren, dass man dem „Feinsten“ das Kommando „groß“ voranstellt.

Die neue Version erfüllt zudem viele alte Wünsche, an erster Stelle der Import von DSS-Pro-Dateien professioneller Diktiergeräte: das ist für Ärzte und Anwälte ein Schritt nach vorn. Um Handmikrofone der Hersteller Philips und Olympus perfekt einzubinden, benötigt man wie gehabt Zusatzsoftware, hier hat vor allem Philips mit seinem „Device Control Center“ gute Arbeit geleistet, es funktioniert auch mit Dragon 11 prima.

Die „Verarbeitungsanzeige“ während der laufenden Erkennung wurde durch ein kleines, sich drehendes Nuance-Symbol ersetzt, man wird also weniger abgelenkt. Deutlich einfacher gerät nun die Verwaltung unterschiedlicher Benutzerprofile, und bei der Korrektur sind ebenfalls zahlreiche Verbesserungen zu beobachten. Richtiges Korrigieren bleibt entscheidend für die Lernfähigkeit von Dragon – und hier wird leider der Einsteiger mit der zwar erweiterten, aber noch immer dürren Online-Hilfe nicht besonders weit kommen. Wer den Drachen perfekt beherrschen will, muss einem Profi über die Schulter gucken.

Jenseits der Erkennung gesprochener Texte hat Dragon Naturally Speaking auch bei der PC-Bedienung zugelegt: „Durchsuche Wikipedia nach Frankfurter Allgemeine Zeitung“ versteht die Software auf Anhieb, aber auch „Durchsuche Facebook nach Max Müller“. Für uns der wichtigste neue Befehl: „Programme auflisten“ zeigt alle geöffneten Windows-Fenster. Wir probierten das Programm

mit einem Core-2-Duo-Prozessor T5250 (1,5 Gigahertz) unter Windows 7 sowie einem T6300 (1,8 Gigahertz) unter Windows XP – das geht gerade noch ordentlich, etwas mehr Rechenleistung kann nicht schaden. Endlich werden mit Version 11 auch Mehrkern-Prozessoren optimal angesprochen.

Die amerikanische Software ist in mehreren Varianten erhältlich: Die teuren juristischen und medizinischen Versionen sollte man ebenso wie Professional (830 Euro) nur im Fachhandel nach entsprechender Beratung kaufen. Die von uns getestete Premium-Variante für 200 Euro ist das Paket der Wahl für die meisten Käufer. Sie unterstützt mehrere Audioquellen an einem Benutzerprofil, ferner das Diktat in so gut wie alle Anwendungsprogramme, auch in Microsoft Word, Excel und Outlook sowie in Open Office Writer. Mit Googlemail funktioniert es nach Angaben des Herstellers ebenfalls, bei uns aber nicht zufriedenstellend. Skeptisch sind wir bei der Einsteigerversion „Home“ für 100 Euro mit stark reduziertem Funktionsumfang. Auch „Premium Mobile“, die für 300 Euro ein billiges Diktiergerät mitbringt, würden wir nicht nehmen.

Mit Dragon Naturally Speaking 11 erfasst der Vielschreiber seine Texte dreimal schneller als mit der Tastatur. Die Ausrede, Spracherkennung sei noch unausgereift, gilt nicht mehr. Die größte Hürde ist das Diktieren selbst, also die im Textverarbeitungszeitalter verlernte Fähigkeit, im Kopf klar zu formulieren.