

# Das Mundwerk der Maschinen



**Sprachsoftware vertont den Brockhaus ebenso wie Webseiten. Die Technik ist so weit, dass die synthetischen Stimmen äußerst menschlich klingen.**

**D**as Brockhaus-Lexikon zum Vorlesen? Das klingt etwas verrückt, doch das gibt es tatsächlich: Wer die neueste Version des „Brockhaus Multimedia Premium“ besitzt, kann sich bequem zurücklehnen und per Kopfhörer mit Wissen berieseln lassen. Die DVD enthält 300 Audiodateien im MP3-Format, die die Lexika-Einträge in gesprochener Form hörbar machen, mehrere tausend weitere können die Nutzer kostenlos aus dem Internet herunterladen. „Es ist mehr als ein zusätzliches Feature“, sagt Produktmanager Jürgen Osterberg. Mit den Audiodateien zielt man auf den Trend des Podcasting ab – Audio und Video werden auf mobile Geräte geladen und später unterwegs angehört und angesehen.

In Sprache umgesetzt wurden die Lexikon-Einträge mit einer speziellen Software. Diese Aufgabe einem professionellen Sprecher anzuvertrauen, hätte vermutlich Monate gedauert, wäre teuer gekommen – und es ist obendrein fraglich, ob ein menschlicher Sprecher die Sisyphos-Aufgabe des Vorlesens von Tausenden Namen und Daten durchgestanden hätte. Statt auf einen menschlichen Sprecher setzte Brockhaus daher auf automatisierte Hilfe, und die synthetische Stimme klingt überraschend menschlich. „Wir hatten bereits Rückmeldungen von Kunden, die sich eine andere Sprecherin wünschten. Sie nahmen wohl an, es handele sich um eine menschliche Stimme“, sagt Jürgen Osterberg. Wurde Sprachsoftware vor wenigen Jahren noch eher belächelt, hat

sie sich inzwischen zur praxisnahen Hilfe gemausert. „Wenn man an sprechende Computer denkt, hat man heute immer noch eine blecherne, kratzende Roboterstimme im Ohr. So etwas aber ist seit Jahren Geschichte. Die Stimmen sind längst sehr viel natürlicher und wirken deshalb sehr viel angenehmer“, erläutert Elisabeth Bauer, Pressesprecherin von Linguattec. Mit der Software des Münchner Unternehmens hat Brockhaus sein Audio-Projekt umgesetzt. Die Software heißt Voice Reader Studio und ist ein professionelles Programm, das nach Angaben von Linguattec bereits ganze Tonstudios ersetzen kann. Ganz so ist es dann zwar nicht, denn obwohl die Stimme – übrigens je nach Wunsch männlich oder weiblich – sehr natürlich klingt und gut zu verstehen ist, ist sie hier und dort abgehackt und der Tonfall ändert sich mit einzelnen Silben.

## Der „Mercedes“ der Sprachsoftware

Trotzdem ist die Software von Linguattec so etwas wie der Mercedes unter der Sprachsoftware, und das äußert sich auch in anderen Produkten wie Voice Pro. Während der Voice Reader im Grunde alles vorliest, was ihm in den Weg kommt – vom Office-Dokument bis zur Webseite –, kann Voice Pro anders herum die Stimme ihres Besitzers in Text umwandeln. Die zugrundeliegende Sprach-Engine der Linguattec-Produkte ist Via Voice, eine Entwicklung von IBM. Big Blue bietet Via Voice auch selbst an, in Tests gefiel das Produkt allerdings nicht so sehr wie das des Partners und Spezialisten Linguattec, der Via Voice um eigene Entwicklungen anreicherte. Andere Software wie Dragon oder Voice Office ist im Vergleich günstiger, hinkt jedoch technisch hinterher. Mehrfach gewann Linguattec Preise für technische Entwicklungen, und immer wieder erntete man auch Wohlwollen in der Gemeinde der Anwender – zum Beispiel, weil man die Sprachtechnologien zum kostenfreien Test ins Internet stellte.

Mit Voice Pro diktieren unter anderem Anwälte oder Ärzte ihrem Computer Briefe und Unterlagen. Die Genauigkeit der Spracherkennung kann hier nach Elisabeth Bauer bei bis zu 99 Prozent liegen. Dazu dürfe man allerdings Aussagen nicht allzu viel Glauben schenken, Software zur Spracherkennung müsse einfach nur installiert werden. „Es ist nicht ganz falsch, dass man ohne langwieriges Sprachtraining gleich loslegen kann. Man hat jedoch mehr Freude und Erfolg mit der Software, wenn man das Sprechertraining absolviert“, so Bauer. Soll heißen: Durch das Training kann sich die Software auf die individuelle Aussprache einstellen oder auch auf den Tonfall. Außer diktieren lassen sich mit Voice Pro auch Programme unter Windows mit gesprochenen Befehlen steuern.

Die hohe Kunst ist allerdings die Verbindung aus Zuhören und intelligenter Reaktion. Solche Visionen von sowohl verstehenden als auch sprechenden Maschinen gibt es, seit es Maschinen gibt (siehe Kasten). Weil die menschliche Sprache aber voller Emotionen steckt, ist eine perfekte Nachahmung durch Computer mit heutiger Technik unmöglich. Immerhin aber ist man der Struktur einzelner Sprachen in Teilen auf die Spur gekommen und kann sie deshalb dem Rechner vermitteln. Während des Zweiten Weltkriegs etwa

fanden englische Wissenschaftler, dass im Deutsch ihrer Gegner das „e“ besonders häufig vorkommt. Auf Basis dieses Wissens konnten die Engländer schließlich verschlüsselte Nachrichten knacken. Ganz ähnlich macht es heute Sprachsoftware. Sie versteht zwar nicht im emotionalen Sinn, erkennt jedoch anhand bestimmter Merkmale, was gesprochen wird. Eine letzte Entwicklung ist das Erkennen von emotionalen Zuständen des Sprechers. Äußerst hilfreich fanden das zum Beispiel Betreiber von Callcentern. Sie filterten mit der Technik etwa laute Stimmen heraus, die zu aufgebracht Kunden gehörten – und bevor sie an menschliche Ansprechpartner gelangten, wurden die Aufgeregten automatisiert beruhigt. Besonders zuverlässig jedoch war diese Erkennung nicht, und man nahm davon wieder Abstand. Ohne Frage aber wird so etwas zur nahen Zukunft automatisierter Callcenter gehören.

## Software per Sprache steuern

Unabdingbar ist Sprachsoftware für Blinde und stark Sehbeeinträchtigte, die mit dem Computer arbeiten. Auf Webseiten sorgt Sprachsoftware für die nötige Barrierefreiheit. Genau wie Rathäuser oder Ämter als öffentlicher Raum über Rampen für Rollstuhlfahrer erreichbar gemacht, so bilden Vorlesefunktionen die Brücke ins Internet. Bei der öffentlichen Hand gehört Sprachsoftware daher zum festen Bestandteil der Webseiten.

Einen guten Schritt weiter ging man beim Straßenbetrieb Straßen.NRW. In der Krefelder Niederlassung arbeitet dank Sprachsoftware eine blinde Mitarbeiterin mit dem komplexen Betriebswirtschaftsprogramm SAP. Die zusätzliche Software liest der Verwaltungsfachangestellten Petra Kautz Beschreibungen und Inhalte der Felder vor, in denen sich der Cursor befindet. „Es war ein sehr schwieriges und aufwändiges Prozedere, die Software an SAP anzupassen“, sagt Petra Kautz. Doch durch die zusätzliche Sprachsoftware namens Supernova des englischen Herstellers Dolphin ist es ihr überhaupt möglich, mit SAP zu arbeiten. Und sich SAP vorlesen zu lassen klingt nicht unbedingt verrückter, als einem Lexikon zu lauschen.

*Dirk Schäfer*

## Zeitreise I Der Traum vom verstehenden Computer

Mitte des 18. Jahrhunderts konstruierte ein österreichisch-ungarischer Mechaniker einen Schach-Apparat in Tischform samt mechanischem Gegner – einer beweglichen Puppe in türkischen Gewändern. Jahrelang fragte sich die Welt, wie dieser „Schach-Türke“ funktionierte und wie es sein konnte, dass er fast immer gewann. Des Rätsels Lösung war ein Kleinwüchsiger, der in dem Apparat steckte.

Versuche, die menschliche Sprache künstlich nachzuahmen, reichen noch weiter zurück: Um die Funktionsweise des menschlichen Sprechapparats zu verstehen, bauten Wissenschaftler im Mittelalter den wohl ersten künstlichen Stimmapparat der Geschichte. Tausend Jahre später suchte man die menschliche Stimme in elektrische Signale umzuwandeln. Die Forschungen aus den 1930er-Jahren sind bis heute Grundlage für die Synthetisierung von Sprache und Tönen. Doch erst die Erfindung des Computers rückte den Traum von einer sprechenden Maschine in greifbare Nähe.