

History of Speech & Voice Recognition and Transcription Software

Dragon Naturally Speaking Medical Transcription Software
Dragon Medical
Dragon Professional
Dragon Preferred
Transcription Software Comparisons
System Requirements
Dragon Hardware
Comp. List
Feedback
Dragon Tutorial
HIPAA Statement
Dragon Add-on Vocabularies
Trigram
HIPAA Made Easy Compliance Software
Main Page
Physicians General
Employer's Medical
Technology History
Speech Recognition
Descriptions
Transcriptions



Please also view the [Speech Recognition Timeline](#)

> SPEECH & VOICE RECOGNITION:

The ability of machines to respond to spoken commands. Speech & Voice recognition enables "hands-free" control of various electronic devices—a particular boon to many disabled persons—and the automatic creation of "print-ready" dictation. Among the earliest applications for speech & voice recognition were automated telephone systems and medical dictation software (Transcription).

"Speech Recognition." Encyclopedia Britannica. 2003. Encyclopedia Britannica Premium Service. 08 Oct. 2003. <<http://www.britannica.com/eb/article/19027>>.

The technology of Automatic Speech (Voice) Recognition (ASR) and Transcription has progressed greatly over the past few years. Ever since research of this technology began in 1936, the largest barriers to the speed and accuracy of speech & voice recognition was computer speed and power (or lack thereof). With the average the CPU now at and above a Pentium III and RAM levels at 500 MB and up, **accuracy levels have reached 95% and better with transcription speeds at over 160 words per minutes.**

As mentioned above, the study of automatic speech recognition and transcription began in the 1930s with ATT&T's Bell Labs. At that time, most research was funded and performed by Universities and the U.S. Government (primarily by the Military and DARPA - [Defense Advanced Research Project Agency](#)). It was not until the early 1980's when the technology reached the commercial market.

Like most emerging technologies, there were several competing research "camps", each working independently to develop speech recognition. Please view the [Speech Recognition Timeline](#) to get a full view of its development.

The first company to launch a commercial product was Covox in 1982. Covox brought digital sound (via The Voice Master, Sound Master and The Speech Thing) to the Commodore 64, Atari 400/800, and finally to the IBM PC in the mid '80s. Along with (or bundled) this introduction of sound to computers came Speech Recognition.

Another company that was founded in 1982 and whose eventual product has become the overwhelming leader in the speech recognition market was **Dragon Systems**. Scansoft, Inc. now owns and manufactures this product, Dragon Naturally Speaking.

[Dragon Systems History](#)

¹ Dragon Systems was founded in 1982 by James and Janet Baker to commercialize speech recognition technology. As graduate students at Rockefeller University in 1970, they became interested in speech recognition while observing waveforms of speech on an oscilloscope. At the time, systems were in place for recognizing a few hundred words of discrete speech, provided the system was trained on the speaker and the speaker paused between words. There were not yet techniques that could sort through naturally spoken sentences. James Baker saw the waveforms—and the problem of natural speech recognition—as an interesting pattern-recognition problem.

Rockefeller had neither experts in speech understanding nor suitable computing power, and so the Bakers moved to Carnegie Mellon University (CMU), a prime contractor for DARPA's Speech Understanding Research program. There they began to work on natural speech recognition capabilities. Their approach differed from that of other speech researchers, most of whom were attempting to recognize spoken language by providing contextual information, such as the speaker's identity, what the speaker knew, and what the speaker might be trying to say, in addition to rules of English. The Bakers' approach was based purely on statistical relationships, such as the probability that any two or three words would appear one after another in spoken English. They created a phonetic dictionary with the sounds of different word groups and then set to work on an algorithm to decipher a string of spoken words based on phonetic sound matches and the probability that someone would speak the words in that order. Their approach soon began outperforming competing systems.

After receiving their doctorates from CMU in 1975, the Bakers joined IBM's T.J. Watson Research Center, one of the only organizations at the time working on large-vocabulary, continuous speech recognition. The Bakers developed a program that could recognize speech from a 1,000-word vocabulary, but it could not do so in real time. Running on an IBM System 370 computer, it took roughly an hour to decode a single spoken sentence. Nevertheless, the Bakers grew impatient with what they saw as IBM's reluctance to develop simpler systems that could be more rapidly put to commercial use. They left in 1979 to join Verbex Voice Systems, a subsidiary of Exxon Enterprises that had built a system for collecting data over the telephone using spoken digits. Less than 3 years later, however, Exxon exited the speech recognition business.

With few alternatives, the Bakers decided to start their own company, Dragon Systems. The company survived its early years through a mix of custom projects, government research contracts, and new products that relied on the more mature discrete speech recognition technology. In 1984, they provided Apricot Computer, a British company, with the first speech recognition capability for a personal computer (PC). It allowed users to open files and run programs using spoken commands. But Apricot folded shortly thereafter. In 1986, Dragon Systems was awarded the first of a series of contracts from DARPA to advance large-vocabulary, speaker-independent continuous speech recognition, and by 1988, Dragon conducted the first public demonstration of a PC-based discrete speech recognition system, boasting an 8,000-word vocabulary.

In 1990, Dragon demonstrated a 5,000-word continuous speech system for PCs and introduced DragonDictate 30K, the first large-vocabulary, speech-to-text system for general-purpose dictation. It allowed control of a PC using voice commands only and found acceptance among the disabled. The system had limited appeal in the broader marketplace because it required users to pause between words. Other federal contracts enabled Dragon to improve its technology. In 1991, Dragon received a contract from DARPA for work on machine-assisted transcription systems, and in 1993, Dragon received a Federal Technology Development Project award to develop, in collaboration with Acoustic Research Corporation, continuous speech & voice recognition systems for doctors.

Dragon Naturally Speaking
Home

Medical Software
Consultant

Dragon Naturally Speaking
Demo

Contact

Order

History of Speech & Voice Recognition and Transcription Software

Dragon Naturally Speaking Medical Transcription Software



Please also view the [Speech Recognition Timeline](#)

Dragon Medical

Dragon Professional

Dragon Preferred

Transcription Software Comparisons

System Requirements

Dragon Hardware Comp. List

Feedback

Dragon Tutorial

HIPAA Statement

> SPEECH & VOICE RECOGNITION:

The ability of machines to respond to spoken commands. Speech & Voice recognition enables “hands-free” control of various electronic devices—a particular boon to many disabled persons—and the automatic creation of “print-ready” dictation. Among the earliest applications for speech & voice recognition were automated telephone systems and medical dictation software (Transcription).

"Speech Recognition." *Encyclopedia Britannica*. 2003. *Encyclopedia Britannica Premium Service*. 08 Oct, 2003 <<http://www.britannica.com/eb/article?eu=139072>>.

The technology of Automatic Speech (Voice)Recognition (ASR) and Transcription has progressed greatly over the past few years. Ever since research of this technology began in 1936, the largest barriers to the speed and accuracy of speech & voice recognition was computer speed and power (or lack there of). With the average the CPU now at and above a Pentium III and RAM levels at 500 MB and up, **accuracy levels have reached 95% and better with transcription speeds at over 160 words per minutes.**

As mentioned above, the study of automatic speech recognition and transcription began in the 1936 with ATT&T's Bell Labs. At that time, most research was funded and performed by Universities and the U.S. Government (primarily by the Military and DARPA - [Defense Advanced Research Project Agency](#)). It was not until the early 1980's when the technology reached the commercial market.

Like most emerging technologies, there were several competing research "camps", each working independently to develop speech recognition. Please view the [Speech Recognition Timeline](#) to get a full view of its development.

The first company to launch a commercial product was Covox in 1982. Covox brought digital sound (via The Voice Master, Sound Master and The Speech Thing) to the Commodore 64, Atari 400/800, and finally to the IBM PC in the mid '80s. Along with (or bundled) this introduction of sound to computers came Speech Recognition.

Another company that was founded in 1982 and whose eventual product has become the overwhelming leader in the speech recognition

Technology History

Speech Recognition



Descriptions

Transcription

Medical Transcription

Voice Recognition

Voice Recognition Software

Dragon Naturally Speaking

market was **Dragon Systems**. Scansoft, Inc. now owns and manufactures this product, Dragon Naturally Speaking.

Dragon Systems History

¹ Dragon Systems was founded in 1982 by James and Janet Baker to commercialize speech recognition technology. As graduate students at Rockefeller University in 1970, they became interested in speech recognition while observing waveforms of speech on an oscilloscope. At the time, systems were in place for recognizing a few hundred words of discrete speech, provided the system was trained on the speaker and the speaker paused between words. There were not yet techniques that could sort through naturally spoken sentences. James Baker saw the waveforms--and the problem of natural speech recognition--as an interesting pattern-recognition problem.

Rockefeller had neither experts in speech understanding nor suitable computing power, and so the Bakers moved to Carnegie Mellon University (CMU), a prime contractor for DARPA's Speech Understanding Research program. There they began to work on natural speech recognition capabilities. Their approach differed from that of other speech researchers, most of whom were attempting to recognize spoken language by providing contextual information, such as the speaker's identity, what the speaker knew, and what the speaker might be trying to say, in addition to rules of English. The Bakers' approach was based purely on statistical relationships, such as the probability that any two or three words would appear one after another in spoken English. They created a phonetic dictionary with the sounds of different word groups and then set to work on an algorithm to decipher a string of spoken words based on phonetic sound matches and the probability that someone would speak the words in that order. Their approach soon began outperforming competing systems.

After receiving their doctorates from CMU in 1975, the Bakers joined IBM's T.J. Watson Research Center, one of the only organizations at the time working on large-vocabulary, continuous speech recognition. The Bakers developed a program that could recognize speech from a 1,000-word vocabulary, but it could not do so in real time. Running on an IBM System 370 computer, it took roughly an hour to decode a single spoken sentence. Nevertheless, the Bakers grew impatient with what they saw as IBM's reluctance to develop simpler systems that could be more rapidly put to commercial use. They left in 1979 to join Verbex Voice Systems, a subsidiary of Exxon Enterprises that had built a system for collecting data over the telephone using spoken digits. Less than 3 years later, however, Exxon exited the speech recognition business.

With few alternatives, the Bakers decided to start their own company, Dragon Systems. The company survived its early years through a mix of custom projects, government research contracts, and new products that relied on the more mature discrete speech recognition technology. In 1984, they provided Apricot Computer, a British company, with the first speech recognition capability for a personal computer (PC). It allowed users to open files and run programs using spoken commands. But Apricot folded shortly thereafter. In 1986, Dragon Systems was awarded the first of a series of contracts from DARPA to advance large-vocabulary, speaker-independent continuous speech recognition, and by 1988, Dragon conducted the first

public demonstration of a PC-based discrete speech recognition system, boasting an 8,000-word vocabulary.

In 1990, Dragon demonstrated a 5,000-word continuous speech system for PCs and introduced DragonDictate 30K, the first large-vocabulary, speech-to-text system for general-purpose dictation. It allowed control of a PC using voice commands only and found acceptance among the disabled. The system had limited appeal in the broader marketplace because it required users to pause between words. Other federal contracts enabled Dragon to improve its technology. In 1991, Dragon received a contract from DARPA for work on machine-assisted translation systems, and in 1993, Dragon received a federal Technology Reinvestment Project award to develop, in collaboration with Analog Devices Corporation, continuous speech & voice recognition systems for desktop and hand-held personal digital assistants (PDAs). Dragon demonstrated PDA speech recognition in the Apple Newton MessagePad 2000 in 1997.

Late in 1993, the Bakers realized that improvements in desktop computers would soon allow continuous voice recognition. They quickly began setting up a new development team to build such a product. To finance the needed expansion of its engineering, marketing, and sales staff, Dragon brokered a deal whereby Seagate Technologies bought 25 percent of Dragon's stock. By July 1997, Dragon had launched Dragon NaturallySpeaking, a continuous speech & voice recognition program for general-purpose use with a vocabulary of 23,000 words. The package won rave reviews and numerous awards. IBM quickly followed suit, offering its own continuous speech recognition program, ViaVoice, in August after a crash development program. By the end of the year, the two companies combined had sold more than 75,000 copies of their software. Other companies, such as Microsoft Corporation and Lucent Technologies, are expected to introduce products in the near future, and analysts expect a \$4 billion worldwide market by 2001.

In 2000, Lernout & Hauspie acquired Dragon Systems. In 2001, Scansoft, Inc. acquired all rights to Lernout & Hauspie's speech recognition products including Dragon Naturally Speaking. In 2003, Scansoft, Inc. acquires Speechworks.

Scansoft, Inc. is presently the world leader in the technology of Speech Recognition in the commercial market.

¹ Funding a Revolution: Government Support for Computing Research . Copyright 1999 by the National Academy of Sciences. http://www.nap.edu/readingroom/books/far/ch9_b2.html

SOURCE: The primary source for this history is Garfinkel (1998).

[About Us](#) [Contact Us](#) [Site Map](#)
[Dragon Naturally Speaking](#) [Medical Transcription](#)
[Voice Recognition](#) [Voice Recognition Software](#) [Transcription](#)