



**The Prognosis is Good:
Speech Recognition Software
Can Increase Productivity
in the Medical Environment**

Introduction

Traditionally viewed as simply a means of dictating text into a personal computer, today's speech recognition software can play a far more significant role in the medical environment. In addition to pure dictation, speech recognition software can be used to manage e-mail, speed up repetitive tasks on the PC, reduce transcription and charting costs, speed-up turnaround, and protect employees from repetitive strain injuries. What's more, the software can be integrated with other medical applications to make those programs more effective and easier to use. Rapid hardware advancements and improvements in the technology itself have increased its utility, accuracy, speed, and ease of use. It has brought the cost of ownership to an affordable level for every medical office, or department within an organization. When properly implemented, speech recognition software can increase productivity for every employee.

Like any technology, the implementation of a speech recognition program should be carefully planned so as to achieve the full benefit of the software and maximize the return on your investment. This white paper provides an overview of the basics of how the software works, what medical offices or departments can do with speech recognition, and recommendations for implementation.

What Is Speech Recognition Software?

Speech recognition software uses the human voice as the main communication mechanism between the user and the computer. While relatively simple to use, speech recognition software is sophisticated technology that uses "language modeling" to recognize and differentiate among the millions of human utterances that make up any language.¹

The software enables users to input text and data into virtually any Microsoft® Windows®-based application by voice, as well as to navigate the computer desktop with little or no use of their hands. Users speak naturally into a noise-canceling microphone connected to the computer.² The software "recognizes" the spoken words, converts them into text, and displays them on the screen.

Most speech recognition programs also allow users to speak a standard command that prompts the computer to perform an action. For example, the user says, "Start WordPerfect," and the PC launches WordPerfect®. The more advanced speech recognition programs also enable users to create customized commands (macros), such as "Send an e-mail to Doug Z," which will open up their e-mail application, with Doug Z's address appearing in the address line.

The enrollment process is simple. After installing the software, each user must read aloud from a choice of prepared texts for about five minutes. Based on the dictation that it captures, the software analyzes how the user pronounces each word and stores the data to prepare a unique user profile, or "User" for that individual.

'How do speech recognition software programs understand speech?

Speech recognition software programs are based on statistical probability. The software analyzes an incoming stream of sounds and interprets those sounds as commands and dictation. This process of interpretation is called speech recognition, and its success is measured by the percentage of correct interpretations, or recognition accuracy.

The software relies on three sources of information to achieve high recognition accuracy:

- **Acoustic model**—a mathematical model of the sound patterns used by the speaker's language.
- **Vocabulary**—a list of words that the program can recognize. Each word in the vocabulary has a text representation and a pronunciation.
- **Language model**—statistical information associated with a vocabulary that describes the likelihood of words and sequences of words occurring in the user's speech.

When you create and train a user, you start with a standard set of models and then customize them for the way you speak (acoustic model) and the way you use words (vocabulary and associated language model). The software employs your customized user files to guess the words that you said.

²The quality and type of noise-canceling microphone is a critical success factor in implementing speech recognition. See the section on hardware recommendations for more information.

As the individual uses the software and corrects recognition errors, the software becomes increasingly accurate. Most programs enable users to add new words or customize the vocabulary for their particular practice or specialty. Using specialty vocabularies can heighten accuracy even further. Some speech recognition software programs include a medical vocabulary—incorporating diseases, medications, procedures, and acronyms in addition to the standard business vocabulary—and can automatically recognize and format prescriptions and patient encounters. For certain programs, specialty medical vocabularies can also be created in-house or purchased from third-party sources.

Users can also dictate into a handheld digital recorder for the user or an assistant to download onto a PC and edit. Instead of transcribing from scratch, the assistant will download the file, listen to the recorded dictation while reading the text on screen, and make corrections or edits. Using speech recognition software to input the text can drastically reduce the cost and turnaround time of transcription.

Uses of Speech Recognition Software in a Medical Environment

Each practice and/or individual uses speech recognition for different purposes, depending on their responsibilities, workflow, preferences, and other applications used. Speech recognition users include doctors, medical students, nurses, physician assistants, nurse practitioners, dentists, physical/occupational therapists, pharmacists, administrators, people with disabilities, transcriptionists, assistants, and other support staff.

Dictation

Dictation is the most versatile and widespread use for speech recognition software. Some individuals can't or prefer not to type, either because they are untrained as typists, have a disability, or wish to prevent the development of a repetitive strain injury (see below). Many practices have decreased the number of support staff and require physicians to generate their own records. Even doctors who typically dictate documents for others to transcribe may use speech recognition occasionally, when they need to produce a document on the spot or after hours, or when they are responding to e-mail.



Most people can talk faster than they can type. For example, a relatively fast typist who can type 50 net words per minute³ can produce a 300-word e-mail in 6 minutes. Using speech recognition software, a person dictating 140-160 words per minute without any errors can produce the same 300-word e-mail in about 2 minutes – one-third of the time. This does not include the additional time the person can save by using voice commands to open the e-mail program, look up an e-mail address from their contact management software programs, and send the e-mail by voice.

Doctors who wish to maintain their traditional workflow can dictate into a handheld recorder⁴ or save their recorded dictation⁵ with their documents for someone else to transcribe or correct. This can substantially reduce the turnaround time over traditional transcription. If transcription is produced in-house, using speech recognition software frees up support staff for more productive tasks. If transcription is outsourced for correction, it can significantly reduce costs.

³Net words per minute are determined by measuring a person's average gross speed in words per minute and subtracting the number of errors made.

⁴The handheld recorder is typically a digital recorder. Not all recorders work with all speech recognition software programs. Check with the software manufacturer to confirm if a recorder is approved for use with their product.

⁵Some speech recognition programs enable users to save their recorded dictation with their text file so that they or a third-party can correct or edit the file while listening to or periodically checking the original dictation. Check with the software manufacturer to confirm if this feature is available.



Navigate the Windows® desktop by voice

Speech recognition software enables users to “command-and-control” the computer desktop by voice. Virtually any menu item or dialog box can be controlled hands-free. Users can edit and format their work, launch applications and open files, cut and paste documents, and insert standard blocks of text or their scanned signature, enabling them to speed up routine tasks on the PC.

Create, manage, and send e-mail

Managing e-mail takes up an increasing amount of everyone’s day. Speech recognition software can be customized so that users can create, navigate, respond, and send e-mail, all by voice, using their e-mail program, such as Lotus Notes® or Microsoft® Outlook®. In addition, some speech recognition programs contain text-to-speech technology that allows users to have their e-mail documents read aloud, which enables them to complete other tasks while listening to their e-mail.

Speed up repetitive tasks

Repetitive tasks, such as data entry or form filling, can be sped up using speech. In many cases, users who are unfamiliar with complex software programs are more comfortable “telling” the computer what to do than trying to master the interface. Macros can be created to enable users to go from field to field by voice, or to perform a sequence of keystrokes or mouse movements. Even skilled typists are often slower at entering numbers than letters. Patient records can be created and edited using voice commands.

Create a paperless office

Many practices seek to convert all paper documents into electronic files to facilitate back-ups and provide remote access to staff or patients. By making it easier for anyone to create, format, dictate into, search, and manage electronic documents by voice, speech recognition software can facilitate the move to a paperless office. Most Windows®-based applications can be navigated by voice using speech recognition software.

Increase productivity on the road

Physicians and others can increase their productivity during travel time or when away from the office by dictating into a handheld recorder (see above) for transcription later. In addition, some software programs enable users to easily export their User file via the network or portable storage device for use on another computer or laptop so that they can use speech recognition anywhere.

Work on the Web by voice

Speech recognition software enables users to search the Web, access information, and navigate Web pages by speaking URLs and links.

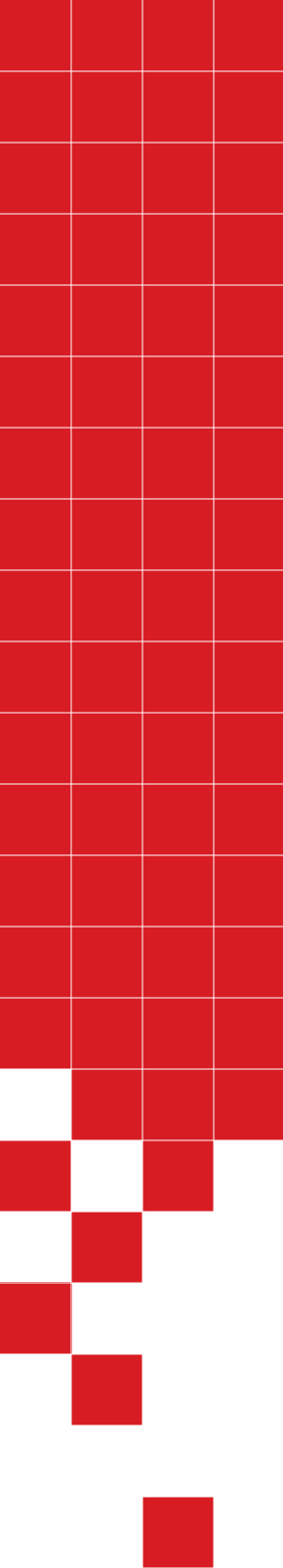
Speech-enable medical applications

Many medical applications can be more effective and easy-to-use when deployed in conjunction with speech recognition. Searches, queries, and form filling are all faster to perform by voice than keyboarding. Charting, prescription writing, aftercare instructions, order entry, database searches, document assembly/automation, and patient records management software programs are all highly conducive to control by speech. Tasks such as text and data entry can be completed by voice in most of the programs without any customization. Other functions can easily be performed using macros or by speech-enabling the application using the software manufacturers’ SDK.

Avoiding repetitive stress injuries

Musculoskeletal disorders (MSDs), including repetitive strain injuries (RSIs), are the single largest job-related injury and illness problem in the United States. According to the Occupational Safety and Health Administration (OSHA), 1.8 million U.S. workers experience work-related disorders annually.⁶

⁶OSHA Fact Sheet, “Ergonomics By the Numbers.”



RSIs, which are often incurred by employees working at computers, are the most common MSD. RSI occurs when muscles or tendons are repeatedly overused or forced into an unnatural position. Keyboarding, clicking, and maneuvering the mouse can strain and damage muscles and tendons in the fingers, hands, wrists, and arms.

The widespread use of computers in the workplace has contributed to the ubiquity of RSI pain and discomfort. OSHA has identified repetition, such as using a keyboard and/or mouse steadily for more than 4 hours daily, as a risk factor that could cause a repetitive strain injury or musculoskeletal disorder. “Intensive computer use accounts for a significant number of MSDs each year, and occupational computer use is growing,” OSHA reports.⁷

While most RSI sufferers are able to find appropriate treatment and return to their positions, some become permanently disabled and are never able to use their hands to operate a computer again. “Workers with severe MSDs often face permanent disability that prevents them from returning to their jobs,” adds OSHA.

Speech recognition software can minimize or eliminate keyboarding and mouse movements that can damage and strain muscles, tendons, and nerves with excessive repetition over time. By giving employees with intensive computer use access to speech recognition software, you can prevent injury before problems arise or help employees return to work sooner, reducing workers’ compensation, medical, and replacement labor costs.

Assisting with ADA Compliance Strategies

Title I of the Americans with Disabilities Act of 1990 (ADA) prohibits employers from discriminating against qualified individuals with disabilities. The workforce includes many qualified individuals with disabilities who can productively use computers when equipped with speech recognition software and supporting hardware and software. Hiring and retaining qualified workers with disabilities is not only a smart employment practice for most employers, it’s the law.

Since speech recognition software can help employers hire and maintain qualified workers with RSI and other disabilities, this technology could play an important role in employers’ ADA compliance strategies.⁸

Implementing a Speech Recognition Software Program

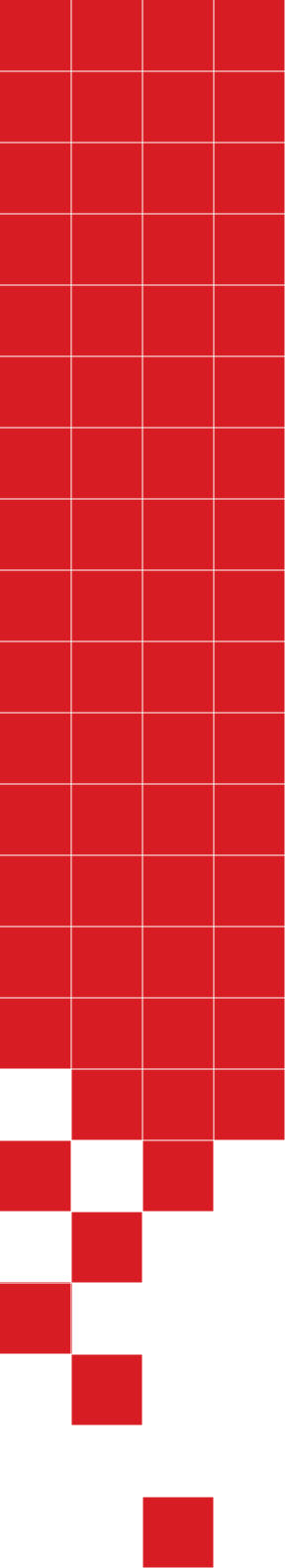
Successful implementation of a speech recognition software program requires careful attention to hardware, user training, and customization. Some practices manage their own speech recognition installation, customization, and training, but most prefer to outsource this work to the software manufacturer, a system integrator, or a speech recognition value added reseller (VAR). An on-site evaluation (pilot) is highly recommended.

Hardware Recommendations

Most practices develop a standard hardware platform for speech recognition users, with alternative options for employees who use speech recognition on a laptop, dictate into a handheld digital recorder, or have special needs. System requirements for speech recognition software vary by software manufacturer. Minimum needs will also vary by the type and number of applications that users deploy. Most speech recognition programs run on PC systems, although some Macintosh®-based products are available.

⁷OSHA Ergonomics Program, Federal Register/Vol. 65, No. 220/Tuesday, Nov. 14, 2000/Rules and Regulations, p. 68343.

⁸The information contained in this white paper does not constitute legal advice. If you have any questions regarding the Americans with Disabilities Act or any other law, you should contact a qualified attorney.



Although speech recognition programs will automatically adjust to the processor and memory of your computer to provide the best combination of accuracy and speed possible, most users will be happier with systems that exceed the software manufacturer's minimum requirements. Speech recognition software is processor-intensive, and, generally speaking, the faster the processor, the better the performance. Users who wish to have multiple applications running at the same time will also benefit from having more RAM on their system.

The computer's sound card is another factor that can affect performance. Speech recognition programs require a soundcard that will accurately process the electrical charges that your voice creates when you speak into the microphone. Static or electrical interference will make it difficult or impossible for users to achieve good speech recognition accuracy. In general, speech programs require a high quality 16-bit sound card. Check with the software manufacturer to verify which sound cards are certified to work with their program.

The software performance can also be impacted by the quality of the microphone used. Speech recognition requires a high-quality, high-level speech signal. Noise-canceling microphones help to block out high ambient noise levels. Most speech recognition programs are sold with a high-quality, noise-canceling headset microphone. Users who do not like wearing a headset may prefer an array microphone; others may opt for a wireless headset. Combined dictation/telephone headsets are also available. Most laptop users achieve high performance with a regular headset microphone, but users who are not able to achieve satisfactory sound quality from their laptops built-in sound hardware may wish to use a USB microphone that processes their voice signal before sending it to the computer. Check with the software manufacturer to verify which microphones are certified to work with its program.

Recommended system requirements for most PC-based speech recognition programs:

- Intel® Pentium® III / 500 MHz processor (or equivalent AMD® processor) or faster
- 256 MB RAM or greater
- 300 MB free hard disk space
- Microsoft® Windows® XP, Millennium, 2000, 98, or Windows NT® 4.0 (with SP-6 or greater)
- Creative® Labs Sound Blaster® 16 or equivalent sound card supporting 16-bit recording
- Microsoft® Internet Explorer 5 or higher
- CD-ROM drive (required for installation)
- Noise-canceling headset microphone
- Speakers (required for playback of recorded speech and text-to-speech features)

User expectations and training

Setting realistic user expectations has a critical impact on the success or failure of a speech recognition program. Although the software itself is easy to install and operate, users who are not accustomed to dictating their thoughts may need practice learning how to dictate. Most physicians who are familiar with dictation will find it easy to adopt speech recognition software. However, they may be used to mumbling or garbling words and expecting the transcriptionist to interpret what they are saying. The quality of the "human sound signal" is just as important as the quality of the sound card.

Although users can begin dictating and using the software after completing their initial five-minute enrollment session, most people increase their productivity when they receive training. Training speeds up the learning curve, instills confidence in users, reduces support costs, promotes the success of a pilot program, and maximizes your investment.



Customization

Users who are just dictating documents for others to transcribe and correct may not need customization, but virtually everyone can benefit from customizing the product to speed up routine tasks. Customization may be as simple as the creation of a macro that inserts your name and title at the end of a letter when you say “my signature” or as complex as a macro that executes a series of keyboard commands and mouse strokes with a spoken command. Macro creation tools are typically included in high-end speech recognition software systems. Although simple macros are easy for users to create, in most cases, firms will achieve better results if a member of IT staff or a speech recognition consultant works with each user to analyze their workflow and customize the program for their needs.

Creating a custom vocabulary including patient, staff, and other physician’s names will increase accuracy. Many speech recognition programs permit custom vocabularies and macros to be exported and shared by multiple users, which decreases the time and cost associated with customization. Individual users can increase their accuracy by running a feature contained in most speech programs that analyzes the user’s written documents to learn their writing style and the words that they use most often.

Conducting a pilot

The majority of practices find it valuable to conduct an on-site evaluation (pilot) with a small number of users before implementing a full-scale speech recognition program. The vendor, or a Value Added Reseller (VAR) can help you set up a pilot, but it is important that you determine your own criteria for evaluating productivity and participant satisfaction before the pilot begins.

For best results, select four to eight computer-savvy employees who want to use speech and are likely to have the time to use the software on a daily basis during the pilot period. A typical pilot, from initial assessment through final evaluation, lasts one to three months. Before the pilot begins, someone from your IT or training department, the vendor, your consultant, or the VAR should sit down with each participant to analyze his or her daily routine. By doing so, custom vocabularies and macros can be developed to enhance their productivity with the software. After the software has been customized for each participant’s needs, group or one-on-one training should be provided.

Conclusion

A growing number of medical practices and physicians have adopted speech recognition software programs to increase their productivity and protect employees from developing repetitive strain injuries. Although implementing a speech recognition program requires careful planning, the resultant cost and time savings can be substantial. For maximum return on your investment, invest in an adequate hardware platform, conduct an on-site evaluation before launching a full-scale program, and incorporate professional training and customization into your program.

The ScanSoft logo and Productivity Without Boundaries are trademarks or registered trademarks of ScanSoft, Inc in the United States and/or other countries. All other products and company names are trademarks or registered trademarks of their respective companies.

Copyright © 2002 ScanSoft, Inc. All Rights Reserved. ScanSoft, Inc. does not provide any assurances on the accuracy of the product. The product can also be subject to changes without notice. ScanSoft, Inc. shall not be responsible for or liable to any customer for decisions made based on this document. The information contained in this white paper does not constitute legal advice.