

Successful Implementation of Speech Recognition Technology

Zephyr-TEC Corp.©2001

Renee L. Griffith, Chief Executive Officer
Zephyr-TEC Corp.
San Mateo, CA

It's time to look seriously at implementing speech recognition software in the workplace. With repetitive stress injuries the leading occupational injury, speech recognition can help reduce and prevent this type of injury. With the recent advances in this technology, it can dramatically increase the productivity of document, spreadsheet and email creation. Speech recognition can assist people with disabilities to access computer technology and participate in the success of a business organization.

This seminar offers information on evaluating various situations for software, training and support programs to successfully implement this technology.

INTRODUCTION

Many factors contribute to the successful implementation of speech recognition. A determination of the tasks or job(s) to be performed using speech recognition is essential to success. Proper training and technical support are required and, particularly for injured or disabled individuals, a reasonable amount of time needs to be planned so the individual has a realistic opportunity to become productive and competitive with the speech recognition software and computer.

SOFTWARE CAPABILITIES

With the variety of speech recognition products available, it can be confusing when trying to decide which package to choose. The latest releases offer the ability to dictate continuously into virtually any Windows application as well as voice activating your desktop.

The high-end packages of continuous speech recognition software can perform many functions, including continuous dictation, allowing the user to input speech directly into various applications, such

as Word or WordPerfect, Excel, etc. in a continuous and natural manner. Some speech recognition software packages allow the user to command and control the Windows desktop and environment including moving between applications and documents, dropping menus, controlling applications, and creating shortcuts.

What it can do:

Navigation is moving through your document to allow editing or formatting. Natural language commands allow you to say "go to the fifth paragraph" instead of driving the cursor to the fifth paragraph using "arrow" navigation commands.

Editing is one of the most productive uses of speech recognition software. Editing commands are used to select text, paragraphs, or pages and cut, copy or paste within the same document, different documents, or even different applications.

Formatting is greatly enhanced with speech recognition. The ability to simply say "justify the paragraph" or "underline the next two lines" is considerably faster than using a mouse to select text and then execute mouse or keystrokes to perform

the formatting. Again, with natural language commands, editing and formatting are frequently combined with continuous natural commands such as "make the next five words red".

Macros and scripts can be developed to voice enable virtually any Windows application and allow hands free use, but it takes time to develop them. With proper training and perseverance, anyone can learn to write macros and scripts that will assist them in their day-to-day operation of computer.

"Mobility" products allow a user to dictate into a digital recording device. The user can then download the speech file from the digital recorder to the speech recognition software--and the words are transformed into a document, unattended. Another individual can then review the document for accuracy, make changes and finalize the document. This basically transforms a transcriptionist into a correctionist.

SPEED, ACCURACY AND VOCABULARY

Recognition rates range from 94% to 98%. The speed of dictation varies according to the type of hardware used and the experience level of the individual. Continuous speech users are enjoying 120 words per minute or more, with 95% or higher accuracy.

An active word vocabulary is one that resides in RAM (random access memory). The total vocabulary refers to the entire available dictionary on your hard disk. These can run up to 300,000 words, with add-on vocabularies and foreign languages available to help accommodate specialized industries and international uses.

Current continuous dictation packages also include special utilities and tools to help build custom vocabularies for industry specific needs. The continuous software "reviews" various documents that you instruct it to process, and produces a word list which you then edit to remove unwanted words or terminology. You will then train any words for

which the speech software doesn't have a language model. Your accuracy should be extremely high when you finish enrollment and vocabulary building.

EQUIPMENT

For optimum performance of today's speech recognition packages, a Pentium III/650 should be considered the minimum. Additionally, a minimum of 128MB RAM is necessary and 256MB of RAM is preferable because today's applications are extremely RAM intensive. All the software packages require a SoundBlaster 16 (or compatible) or higher sound card, but we would recommend using PCI sound card. Each speech recognition package operates with different operating systems including Windows 95, 98, ME, 2000, XP or Windows NT 4.0. All the continuous packages can operate in NT 4.0, however, it is not a simple task to install the speech recognition packages in NT. You must also install at least Service Pack 6a. All the high-end speech recognition software packages ship with a premium headset microphone.

PROS AND CONS

Certainly hands-free operation of a PC is one of the major benefits of speech recognition software. This type of software can potentially reduce chances of Repetitive Stress Injury, (RSI), or give relief to someone already suffering from RSI. For those who do not suffer from RSI, an increase in productivity is possible due to the formatting, editing and navigation controls available. For those who don't know how to type or type slowly, the continuous speech assists in more productive and accurate typing. Finally, speech recognition allows access to jobs previously denied the disabled who can not use their hands to operate a computer.

Voice strain is a potential problem when using speech recognition, especially when used by an individual with RSI or hand disabilities. They will

be using their voice much more than an individual who can use mouse and keyboard combined with speech recognition. Attention should be paid to the voice of the individual to identify any possible problems with voice strain and if necessary an appointment made with a technical voice instructor to learn proper breathing and speaking techniques. Continuous speech recognition software has helped relieve this problem.

MOTIVATION AND PATIENCE

Remember it takes tremendous motivation and patience to be successful with speech recognition software. Injured individuals may take between 30 to 50 hours of working with the software to establish a good voice profile and create the necessary macros to reach productivity. It also takes the software time to formulate how an individual speaks. This is the most frustrating part. Constant attention and stringent correction are required to get excellent recognition accuracy and speed. If the individual doesn't invest this preliminary time, he/she may be frustrated by a low accuracy rate. Even in continuous speech packages with people who are not injured, the initial investment of time will be somewhere around 12 hours to establish a comprehensive vocabulary and high accuracy.

EVALUATION

An evaluation process involves reviewing the applications accessed by the individual and the time percentages spent in each application. Not all jobs requiring the use of a computer can be competitively performed using speech recognition software. Equipment on site should be examined to determine if an upgrade is required or replacement is necessary. After reviewing the particular needs of the individual, a determination can be made regarding which speech recognition software package to recommend. Different packages have different advantages and disadvantages depending on what specific tasks need to be accomplished. Training and

support recommendations should be made only after completing an evaluation.

The environment in which the speech recognition software will be used must be evaluated to ensure there will be no problems with background noise or coworkers. In addition to evaluating speech recognition software and the computer processing power, don't forget to take a serious look at the job where speech recognition will be used. Consider these factors in deciding if an individual can become more productive with speech recognition:

Physical duties

It is necessary to thoroughly identify the physical duties that an individual is required to perform. Speech recognition software can do nothing to relieve physical discomfort from physical activities and it is important to consider that from the beginning.

Computer duties

A complete evaluation/description of the necessary computer duties is required in order to make an accurate assessment as to which speech recognition package is best and how much time will be required to effectively implement the speech recognition software.

Existing application/computer skills of individual

Many times, an individual's assessment of his/her skills can be over- or understated. Often, the individual, supervisor or vocational rehabilitation counselor have different ideas of just exactly how computer "literate" an individual is. Simple software functionality testing can identify the weak areas and training can then be recommended.

Equipment and software

It is vital to assess the current computer equipment used by the individual to determine if an upgrade or replacement of the system is necessary. In addition, it is extremely important to identify the type of operating system utilized (Windows 9X, 2000, XP, NT, network, mainframe or Unix).

Certain speech recognition packages work better with different operating systems. It must be determined if the individual uses standard Windows software packages or whether custom software packages are used. This will influence the decision of which speech recognition package is appropriate. Custom software packages will require integration of the speech recognition software, using macros and scripts.

Environment

Speech recognition software today is extremely robust in its ability to filter out background noise. Offices and cubicles are ideal environments for speech recognition software. Location of the individual's job area is important as well (near noisy equipment, next to the conference room). Surrounding materials in the area, such as glass (large windows) or concrete walls, also make a difference in performance of speech recognition software. Sound "bounces" off these kinds of materials and creates an echo to the microphone. Some adjustments to the environment may be required. All these factors can be addressed in an evaluation.

We offer free phone evaluations upon completion of a Zephyr-TEC Evaluation Form. We address all the above issues and can determine whether or not an on-site Speech Feasibility Study is required. A Speech Feasibility Study includes installation of the software in the working environment and testing with existing applications to determine which software package will integrate best with custom or mainframe environments.

TRAINING AND WHAT IT TAKES

For those who have little or no computer experience, it is absolutely imperative that they learn how to operate a computer using speech recognition from the beginning. This would apply if an individual had performed a job function not

involving a computer but is being retrained to a computer position. This individual will require classes in speech recognition, Windows and any other software package required to perform a specific job or computer task. Even the experienced computer user, familiar and comfortable operating a PC in the Windows environment, should take a minimum of eight hours of speech recognition training and an additional eight to take complete advantage of the advanced features of the professional editions.

The first eight hours consist of an overview of the speech recognition software, proper voice dictation techniques, program commands, simple voice macros and establishing a solid and accurate voice profile. This training can take place at an approved, certified training facility, or in the actual work environment. It is quite often taught in a group setting on-site. In either case, students should return to their computers to practice what was learned prior to completing the advanced course.

This training also provides development of voice macros to be used in a specific work situation. This is essential for both injured and non-injured individuals and normally takes place at the job site.

INTEGRATION OF EXISTING APPLICATIONS

Just knowing how to use the speech recognition software is half the battle. Knowing how to integrate speech recognition with existing applications is the other half. After successful completion of the first eight hours of speech recognition training, many individuals need help with integration into existing applications. An average of eight to twelve hours of on-site integration will usually complete an individual's training program. If other individuals will be performing the same job or tasks using speech recognition, many macros and scripts that are written can be exchanged with other individuals through a simple import/export process. This

reduces the integration costs for subsequent individuals.

The experienced speech recognition instructor can assist the individual in creating macros and scripts that will help the individual become productive more quickly and suffer less frustration. This is essential to success in areas where a mainframe (call center) or Unix based systems are utilized. Students should be taught to develop their own macros and scripts so they can be independent and continue to grow in their positions with speech recognition software.

TECHNICAL SUPPORT

The speech recognition manufacturers offer a variety of technical support options. Most of them offer at least 30 days of free technical support from the date of purchase. After the initial period, individuals can purchase support contracts directly from the manufacturers.

In addition to this technical support, many speech resellers offer custom support packages. Most of the problems associated with implementing speech recognition can be resolved early on, particularly if a professional evaluation has been performed and proper training received.

Since none of the software packages are 100% bug free, and quite often an individual is working in a

custom mainframe or software application, it is imperative that some form of technical support be considered. The speech recognition manufacturers do not provide support for mainframe, Unix or custom software applications or environments. Training your MIS/IT department is important if more than a few individuals are using this type of software.

CONCLUSION

This technology can be used to help return injured workers to their computer jobs, potentially reduce the risk of RSI, allow access to a computer for the disabled or technology displaced individual and increase productivity for the non-injured worker. It can also be used to enhance the learning experience, especially for the developmentally disabled.

Speech recognition is not just for the disabled anymore! Properly implementing this technology on a wide scale basis can increase productivity, reduce costs involved in the prevention of repetitive stress related injuries and gives computer access to the disabled individual.

Renee Griffith, Chief Technical Officer and Founder of Zephyr-TEC Corp. has established a national reputation in her field and has been a featured speaker at national and local conventions of organizations involved in ergonomics, safety and risk management. After becoming disabled in 1991 with DeQuervaine's disease, she utilized speech recognition software to start Zephyr-TEC, now an industry leader specializing in speech recognition training, implementation, integration and support.

Corporate Headquarters:
Zephyr-TEC Corp.
1660 S. Amphlett Blvd.
Suite #116
San Mateo, CA 94402

Toll-free: 800-520-5373
or 650-655-2405
Fax: 650-655-2406

Southern California:
Zephyr-TEC Corp.
9267 Haven Ave. # 260
Rancho Cucamonga, CA 91730

Toll-free: 877-493-7497
or 909-944-8549
Fax: 909-484-1610

Email: info@zephyr-tec.com

Web: www.zephyr-tec.com