

Voice recognition

The area of voice recognition is one of the most exciting areas of computer access and also one of the most misunderstood. The following quote from Tyler Carpenter, an engineer at Dragon Systems, says it all about this area of computer access:

Like all technology, voice recognition is not a panacea, and those who believe it to be "the answer" for everybody (or even anybody) without taking a hard look at its limitations are being just as foolish as those who dismiss it without considering its potential benefits.

Voice recognition has long been termed "the next big thing" and with the availability of better, faster computers, its time may finally be here for many of the consumers with disabilities we work with.

When we discuss voice recognition software, there are two types: discrete and continuous. The original software packages in this area were all *discrete* systems. The defining feature of these discrete packages was that users needed to speak one word at a time, with a slight pause between words. The computer would then try to recognize each word individually. For some users it was very hard to speak in this manner and they would lose their thought. For other users, especially users with speech impairments, this system worked very well. Since the computer was listening to each word individually, as long as the pronunciation of each word was consistent the system could be trained to recognize that word.

The software packages that are available from manufacturers today are considered *continuous* speech products. With these products, the individual speaks in a normal, conversational rate. Instead of listening to individual words, this system listens for words in phrases and in context. This, along with increased processor speeds, produces almost real time recognition of spoken words.

Features of voice recognition software

While there are a variety of voice recognition software packages available today, they all have some similar features. Each product, as part of the *enrollment process*, will require the user to "train" their voice file. To do this, the user must read aloud from a passage provided by the software package. This first step can be a very large obstacle for some individuals who struggle with reading. There have been some strategies developed to overcome this obstacle for users with learning disabilities and visual impairments, who would like to use this product but cannot see the enrollment screens on the monitor. Team members have paused the voice training to read the next passage to the user, then turned on the microphone and had the user speak in the information. This will obviously increase the time needed to complete the enrollment, but this will get certain users over the reading/visual obstacle.

The packages will also differ in the vocabulary size and the ability to provide command and control over the computer environments. To make this a "hands-free" application for users, they must have a package that will accept commands to start programs, open and close windows, and control the mouse. Some of the higher end packages also provide a "Macro" feature, which allows the user to execute a string of commands and/or text to increase input speed.

Other features the service provider needs to be aware of include the ability of the software to function in a network environment and with standard computer applications (such as word processing, database software, etc.), as well as the ability to work with other assistive technology products (such as screen readers). There are products available to enable speech products to work with AT software applications and there also may be after-market products available for custom business applications.

System requirements

As with any software program you purchase, voice recognition products come with a list of system requirements that are suggested by the manufacturer. These requirements are typically "just enough" to allow the program to run. To give the user the best chance of having success with this technology, you should take the manufacturer requirements and double them, at least! There are other strategies you can use to increase computer performance. When the user is using voice recognition software to dictate into a large word processing program, like Microsoft Word, the computer resources will be drained. One suggestion for increased system stability is to use voice recognition software to dictate into the included word processor (such as Dragon Naturally Speaking Dragon Pad) or into a smaller word processor program like WordPad in Windows. Then when the person is done, they simply cut and paste their text into another word processor for editing and final touches.

In years past, there was debate over using voice recognition on a laptop versus a desktop computer. Since laptops have so many parts crammed into a small space, there was sometimes too much noise from internal components and the recognition suffered. Sound cards were also a huge issue, where certain sound cards were better for recognizing speech than others. There are still some components that are better than others and if you visit the voice recognition manufacturer web sites, many will have listings of approved components that have been tested and work well with their products.

Factors to consider

As we continue to discuss voice recognition, there are many additional factors we must consider before determining the best match between the consumer and this technology.

Microphone

This could be the single most important piece of equipment in this technology solution. A good microphone, which meets the user's needs and produces good quality sound into the computer, will reduce the frustration level for the consumer and make this a good technology match.

Many of the software packages will come with a pre-packaged microphone. These are not typically the best microphone on the market. Common sense tells us that if I pay \$125 for a software package, the included microphone will not be worth \$100! The microphones are low quality and we typically encourage teams to spend the extra money and purchase a quality after-market microphone that will serve the user well.

There are many types of microphones available, each with its pros and cons.



[The picture above shows a standard microphone, which plugs into the sound card of the computer with a 1/8" plug]

Standard microphone with 1/8" plug - these are perhaps the most common type of microphone and the type that would typically be packaged with the voice recognition package. These microphones plug directly into the

sound card of the computer and send analog signals. These analog signals are then converted in the computer to digital signals.



*Universal Serial Bus Digital Audio Interface
No need for a sound card*

[The picture above shows a headset microphone that connects to the computer via the USB plug]

USB microphone - instead of connecting to the computer through the sound card, these microphones connect via the USB connector. The benefit of this type of microphone is that the signals coming from the microphone are already converted into a digital signal before it gets to the computer.



[The picture above shows a desk top directional microphone.]

Directional microphone - For a user who can't wear a headset microphone, this type of microphone may be a good solution. These desk top models will allow the user to come and go from the computer without removing the headset. This is also a good type for a user who accesses the computer from a wheelchair. One issue with these types of microphones is that they only register sound in a very narrow field and if the user moves slightly out of that range, they may lose recognition quality. Also, if there is noise behind the person and still in the pick up range, quality will suffer.



[The picture above shows a desktop array microphone.]

Desktop array microphone - An array microphone may look similar to the directional microphone above but it functions much differently. Instead of having one microphone, there are many microphones inside this device, some have as many as seven. The device registers sound from all the microphones and then finds the strongest signal (which is hopefully the user's voice!) and uses this as the input to the computer. All the other signals are cancelled out so they don't interfere. This is a good option for someone who needs a microphone on the desk top but may not be in the exact same position all the time.



[The picture above shows an earset microphone, which contains both a microphone and a speaker.]

Earset microphone - This type of microphone is dual function - the user places it in their ear and the device contains both a microphone and a speaker. These devices are very popular with cell phones right now.



[The picture above shows a portable digital voice recorder.]

Digital voice recorders - With this option, users can dictate from multiple locations and then go to their computer to transfer the voice files into the voice recognition program. While this gives the user many options for dictation locations, it also means that there may be varying degrees of background noise for the computer to try and listen to. This could lead to poor speech recognition.

The placement of the microphone is also important in the sound quality. If the person is wearing the microphone on their head, it is important to place the microphone approximately one quarter of an inch from the corner of their mouth. The microphone should never be directly in front of the mouth. This will cause the microphone to pick up all the "pops" and "clicks" in our speech, as well as the sound of our breathing. If these sounds are picked up by the microphone, the system will try to interpret these sounds as speech and the user will get bad speech recognition. The pictures below illustrate the placement of the microphone.