



Computer Prompting & Captioning Co.

1010 Rockville Pike, Suite 306
Rockville, Maryland 20852-1419, USA
301-738-8487 • FAX 8488 • TTY 8489
info@cpcweb.com • www.cpcweb.com

CPC-500 VoiceCaptioner™ **Speech Recognition Based Realtime Captioning**

The CPC-500 VoiceCaptioner™ does realtime closed captioning using speech recognition software.

It is easy to learn and use, and provides a cost-effective means of captioning live presentations.

The distinguishing feature, which makes it so cost-effective, is its use of speech recognition software to transform spoken words into a text file. This software takes the place of a highly skilled and highly paid realtime court reporter.

The first important question usually asked about the system is: how accurate is it?

The answer is approximately 90% to 95% accurate after training the speech recognition software to the person's speech.

The second question is usually, is this level of accuracy high enough to satisfy FCC, ADA and Section 508 regulations? We believe the answer is yes. However, in today's litigious society, in which a company can be sued for selling coffee, which is too hot, we recommend that you seek advice from your legal department for a definitive answer to this question.

People also want to know, can the software recognize the speech of a number of people, such as the rapid banter back and fourth between newscasters. The answer is no.

The software can only recognize the speech of the person whose speech model is currently loaded into the software. It takes about 30 seconds to load a new person's speech model into the software, and many, many speech models can reside in the computer's memory.

So, how can the software be used to caption a program in which there are many people speaking? The answer is through an off-camera person who previously trained the speech recognition software to their speech, and who repeats what the on-camera people are saying. This person is known in the industry as a "shadow speaker."

It takes a while to develop the ability to act as a shadow speaker, but it is very possible to do so. For example, at the United Nations, simultaneous translators not only repeat what diplomats say, but also simultaneously translate their speech into another language. The typical shadow speaker only repeats what the talent is saying, and does not have to simultaneously translate it into another language.

Speaking of languages, the speech recognition software has both English and Spanish speech recognition engines as of May 2005.

Now let's talk about how the system actually works.

A person speaks into a microphone attached to a computer. Their speech is transformed into a text file by speech recognition software, such as ViaVoice, running on the computer.

The text file is in turn formatted into a closed caption file by the CPC-500 software running on the same computer. This closed caption file is output to a standard closed caption encoder via a serial cable.

Video of the program is simultaneously input into the closed caption encoder through a standard video cable. The closed caption encoder encodes the caption text onto line 21 of the vertical blanking interval. The output of the encoder, which is standard video with embedded closed captions, is then broadcast out and/or recorded for later use.

There is typically a delay of about 3 seconds from the time a person says a word, until that word is output as a closed caption. This is very close to the delay time when a court reporter is used.

Now, how long does it take to train the speech recognition software to attain 90% to 95% accuracy? The answer varies from person to person.

A related, and more easily answered question is: what type of training will achieve this accuracy level. The answer is basically two types.

The initial training typically consists of two to three hours of reading a pre-written text file. The software comes with a several text files to read. However we recommend that you import a text file consisting of the type of dialogue that is typical of what you will be speaking on the programs you will use the VoiceCaptioner for.

For example, if you will be using it for the weather segment of your program, have the weatherperson read the text of a number of previously broadcast weather reports. This is recommended because the speech recognition software will learn to recognize the vocabulary most frequently used for the program.

The follow up training, which will be used to increase the accuracy from its initial level, consists of re-training the software to recognize the words, which it incorrectly recognized during the initial training. This is a key requirement of the training, and you probably will not achieve 90% to 95% accuracy if you do not do this.

After each use of the software, run a spell check to find out which words were incorrectly recognized by the software. Then enter the correct spelling of each word the software recognized incorrectly. Finally, read each of these words into the software. In this manner, the software will learn how to correctly recognize these words.

Through progressive re-training sessions, the software's accuracy will slowly increase. This will not happen overnight, but it does work. It requires a commitment from you to do this re-training. The reward will be increased accuracy as you do the retraining.

After a while, you will reach a level where re-training is rarely needed. At this point, you typically will only have to train the software to learn how to spell the names of people and the names of cities, and other terms not typically used previously by you.

The speech recognition software CPC currently recommends for use is ViaVoice. This software will most likely reach a 90% to 95% accuracy level after training. This is good, but not great. If you require higher levels of accuracy, we recommend that you use a realtime court reporter and not speech recognition software.

Since the CPC-500 software can interface to most Windows based speech recognition software, if and when you or CPC find speech recognition software that is superior to ViaVoice, there is a good chance that CPC will build a interface to allow you to use it.

If you have any questions, feel free to contact CPC at 1-800-977-6678 or email us at sales@cpcweb.com. Also, for more information visit our web site www.cpcweb.com/Captioning/cap_500.htm