

SpeechScape™

Specialized Speech Recognition Vocabularies

A vocabulary is a database of known phonemes, words, and phrases for the speech engine. The active vocabulary is the memory-based part of the full-vocabulary. It is stored in RAM for fast access. The backup dictionary is some 250,000 words or more (plus their phonetic spellings and language modeling) stored on the hard disk. A constrained grammar consists of built-in voice commands. Speech files represent a combination of acoustic, language, and other information about a user's dictation. An acoustic model reflects the way the user pronounces. A language model reflects what words a speaker uses and in what combination with others. It consists of a statistical model regarding how often a word is used by itself and in combination with others. Each individual will have a unique acoustic and language model. This is because every person pronounces words differently and uses them slightly different than any other. Vocabularies and voice commands may be shared. As the speech recognition software is trained, the speech files and vocabulary may be modified.

The backup dictionary in a speech engine includes all the non-custom words in the active vocabulary, plus many more. It is usually accessed by the speech engine when using a correction window or adding words to the system. The active vocabulary remains loaded in RAM. The size of the active vocabulary will vary upon the version and may be limited to 64,000 words or less. Since there is a limit how many words can be active at once, a least-used word is moved to the backup dictionary. Most speech recognition systems keep track of which words are used most often, swapping words between active and backup vocabularies automatically. The correction process will move words from the backup dictionary into the active vocabulary if the word is not currently in the active vocabulary. A user who adds new words from his or her documents may add words to the base dictionary. Statistical information added about word usage and context will affect the language model. Addition of commercially prepared topics and vocabularies has the same effect.

In general, accuracy improves as a user adds specialty vocabularies (e.g., legal, medical, insurance, law enforcement) and contextual information to the base vocabulary and language model. Accuracy can be further improved by personalizing the dictionary. For example, a physician can personalize the vocabulary by adding personal words and phrases, referring physicians, local hospitals and clinics, patient names, abbreviations, and so on. This can be done by adding lists of words and analyzing documents. Some programs may limit the number of words that can be added at any one time. Most commercial speech recognition programs permit the user to add new words with context information. Some programs limit the megabyte size of files that can be added in one session. The advantage of this approach is that the speech engine creates a contextual or statistical model, regarding the words or groups of words associated with any particular word. When that word is used, the engine knows what words are likely to be associated with it, thereby increasing accuracy. In summary, you can use SpeechScape™ to create a specialty vocabulary and language model by loading the application. Further personalize your dictionary by using the speech recognition software's tools to analyze your documents, letters, notes, and reports for additional new words and contextual information.

Revised 01/14/02. Price, availability, specifications, and terms are subject to change without notice. Custom Speech USA, Inc. trademarks are indicated. Other marks are the property of their respective owners.

We teach computers how to listen®

Custom Speech USA, Inc.

3 North Court Street, Suite B365

Crown Point, IN 46307

Tel: 219-662-3800 Fax: 219-662-3877

Web: www.customspeechusa.com Email: info@customspeechusa.com