

Language models

Specialized Vocabularies

for Professional Speech By Gabriel F. Groner, PhD

Continuous speech recognition software needs to deal with information specific to the language and subject matter in order to be useful. The software must “understand” whether the user is speaking English, Spanish or German and if their topic is business, sports or medicine.

Accumulating this information typically requires an *acoustic model*, a *vocabulary*, and the *language model*.

- The *acoustic model*, describes the sounds spoken in the specific language.
- The *vocabulary* is a long list of words used for the particular subject matter, with their spellings and pronunciations.
- The *language model* consists of statistical information about usage of each word alone and with other words. For example, the language model may include the number of times that “right,” “Wright,” “turn right,” “right turn,” “right hand,” and “Mr. Wright” occur in a large body of text.

Different speech recognition vendors call the vocabulary together with the language model, that is the subject matter-specific information, a *vocabulary*, *topic* or *context*. For simplicity, this article uses the term *vocabulary*.

Most continuous speech recognition software includes a large, general-purpose, vocabulary. Although this may suffice for general or business correspondence, most users find that they need additional

proper names and words specific to their subject matter.

The general user can create his own customized or extended vocabulary, either as he dictates, or by processing representative text that includes his words. This takes time and effort, but can improve recognition performance when dictating letters and reports that refer to people and places, and interests and activities such as cooking and sports.

However, professionals who use a large number of specialized words need something better. To create their own customized vocabularies, either they must dictate for many hours with poor recognition accuracy until the software “learns” the vocabulary, or build a custom vocabulary by processing dictated text.

To build a vocabulary, they must collect a great deal of representative text in computer-readable form, and make sure that spelling, capitalization, abbreviations, acronyms, and so forth are correct. Then, using vocabulary- building software, they must select and pronounce correctly the words they want to add to their vocabulary.

It is possible to get better results faster and more economically by purchasing a commercially available, high quality specialized vocabulary, then customizing it by adding a modest number of proper names and specialized words.

How Vocabularies Are Made

Like other handcrafted products, the quality of speech recognition vocabularies depends on the skill and care of the creators. Professional teams of linguists and transcriptionists build some

vocabularies, whereas small groups of programmers and data processors build others.

The most experienced vocabulary developers are the speech recognition vendors and such companies as Breveon, SpeechLaw and Voice Input Technologies.

To create a specialized vocabulary, the team must:

- Collect the data. A generally useful, quality vocabulary is based on millions of words in thousands of representative letters and reports dictated by hundreds of individuals at several locations throughout the geographical region of interest. For example, a specialized vocabulary for U.S. cardiologists require collecting thousands of progress notes, history and physical reports, discharge summaries, cardiac catheterization reports, disability evaluations, and referral and consultation letters dictated by hundreds of cardiologist nationwide.
- Organized the data. To create vocabularies with appropriate content, the data must be organized by subject matter and type of report.
- Clean the data. To avoid introducing errors of spelling, capitalization, hyphenation, or abbreviation in the user’s recognized

Continued on back>

texts, all such errors must be removed from the data used to create the vocabulary. This is a time-consuming process that should be performed by personnel familiar with the subject matter and the style conventions followed by professionals trained to transcribe that subject matter.

- Compiled the word list and enter word pronunciations. Speech recognition vendors provide software that analyzes the text data to compile the word list and generates initial pronunciations for the words. Linguists, or others familiar with the words, how they are pronounced and their phonetic representations, should check and edit the initial pronunciations.
- Compute word usage statistics from the data and produce a vocabulary for use with a specific speech recognition product. This is usually done with software provided by the speech recognition vendor.

A specialized vocabulary must meet its user's full terminology needs. For example, a vocabulary for cardiologists must support dictating reports, including anatomical and procedural words. However, much of what a cardiologist dictates in discharge summaries and other reports is outside his specialty, so a cardiology vocabulary should also include a full internal medicine vocabulary. The same concept applies to all but the most general subject matters.

To assist its user in producing professional quality letters and

reports, a specialized vocabulary must support the user's style. When medical transcriptionists take dictation, they ensure that capitalization, abbreviations, hyphenation, acronyms, and eponyms conform to their professional organizations style guidelines. A speech recognition vocabulary should follow the same guidelines.

Also, a set of related vocabulary should consistently follow the same guidelines. For example, an internist, cardiologist, and general surgeon in the same practice will want to produce stylistically consistent reports and letters.

How to Evaluate a Vocabulary

Quality measures how well the vocabulary meets the user's needs. First check the supplier's specification and claims. A cardiologist, for example, should not buy a vocabulary that is not intended specifically for cardiologists, that does not claim to support the dictation of cardiac catheterization and EKG interpretation reports, nor that is not built on a strong internal medicine base.

Test the vocabulary to determine if it actually provides good word coverage for the subject matter and includes a reasonable number of proper names. Read specialty words and proper names taken from the user's own dictated reports and letters.

If the user's prior dictations are in computer-readable form, and the speech recognition software includes a vocabulary builder, process dictated reports and letters with the vocabulary builder to determine how many of the user's words are not in the vocabulary.

Using the vocabulary, dictate reports and then examine the recognized text to see if it conforms to professional style guidelines. The test reports should include a wide variety of acronyms, abbreviations,

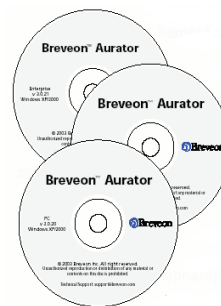
hyphenated words, and units of measure. Medical reports should include both generic and brand drug names. A trained medical or legal transcriptionist should examine the recognized text because there are often many style considerations, some of which are subtle.

Alternately, the user or professional colleague may examine the text to judge its professional appearance and consistency.

A vocabulary built by using many different words, but without sufficient examples of each in context to create a strong language model of word usage, will result in inferior recognition accuracy. The best way to discover this problem is to dictate a wide variety of representative letters and reports.

Continuous speech recognition software for dictation is now both practical and economical. However, as it comes out of the box, this software is not efficient for professional users such as physicians and lawyers. Specialized vocabularies that support the vocabulary needs of such users are commercially available.

Gabriel F. Groner is a consultant in speech recognition technologies and applications and chief technology officer of Breveon Incorporated. He can be contacted at Breveon 1-866-273-8366 or gabe@breveon.com



"After two months of using the Aurator product, I must say that I am most impressed. Prior to using the Aurator; the turnaround time was 3-4 days. Now, it's immediate! This has saved a great deal of both time and money!"

Dr J Shah
Ann Arbor, MI