Using Speech

Speech Recognition Software Still Needs Refinement

By Robin Springer

Speech recognition software has become as easy to find as candy canes at Christmas. From office supply stores to the Internet to consultants who specialize in speech implementation, it's pretty easy to buy the software. But once a user decides he wants to be a dictator, he needs to choose an implementation strategy.

Does he go it alone, guessing which software he needs based on descriptions like Standard, Preferred, Professional, and Premium? After he purchases the software does he forge ahead with little more than a user manual or does he hire someone who will teach him the language of speech recognition and customize the product to interface effectively with the other software he uses?

"In their zeal to sell, manufacturers minimize the importance of training and customization, contributing to end-users' unrealistic expectations," says Ed Rosenthal, president and CEO of Next Generation Technologies.

Attorney Jim Banks agrees, believing his chances of success using Dragon NaturallySpeaking® without hiring a consultant were "zero." Banks says, "There was no way I could have understood the subtleties of the language without sitting down with a pro. You can't get that from the book."

Whether learning from a book or practicing with a professional, anyone who has ever had a puppy knows that however adorable it may be when you first bring it home, if you want it to be well behaved you have to train it, which takes time and commitment. Raising a well-behaved speech profile also requires work.

For the user who just wants to put words on the screen, dictation using an out-of-the-box program may be adequate. But for the person using desktop dictation in a professional environment as a productivity tool, command and control is often also required, which is where customization comes in.

Customizing speech recognition software serves several purposes:

• Increases ease of use

If implementing a task on the computer takes five mouse clicks, but using speech recognition with customized commands a user can complete the task with a single voice command, the user is able to navigate the application faster and more accurately using speech.

And while a professional, whose time is most efficiently spent in her field of expertise, may use voice commands to insert boilerplate text, she probably wouldn't bother with speech if she had to dictate repetitive information into multiple documents.

• Replaces functionality

Speech recognition eliminates the middleman (the keyboard or mouse). While the benefits for people with physical disabilities are often discussed, the software also enables individuals with other types of disabilities. Those with some brain injuries may not be able to type on the keyboard because when they begin concentrating on hitting the correct keys they lose their ability to remember what they want to type. Clients with dyslexia may know exactly what they want to write, but may be unable to process the characters they see on the screen. In both scenarios speech recognition avoids the bottleneck at the mouse and keyboard.

• Introduces new functionality

Customization enables speech recognition to work fully with other programs. For example, the screen reader Jaws, commonly used by people who are blind, does not interface smoothly with Dragon NaturallySpeaking® but the two programs work well together when used in conjunction with software like JSay or JawBone. Similarly, MagnaTalk acts as a bridge between ZoomText, a screen magnifier/screen reader software, and Dragon.

"Out of the box recognition is quite good now," says Rosenthal, "but computer work is not generic. We need to customize speech recognition to increase the degree of efficiency."

VoicePerfect, which manufactures PathSpeak, a product marketed to pathologists, tested the efficiency of speech recognition with varying degrees of customization. Transcribing a digitally recorded 165-word pathology report, each time with the same speech file, they documented the following results (see fig. 1):

NaturallySpeaking® *with PathSpeak*
Errors	1
Clinically Significant	0
Recognition Rate	99%

NaturallySpeaking® *only*
Errors	27
Clinically Significant	22
Recognition Rate	84%

NaturallySpeaking® *with pathology*
Errors	22
Clinically Significant	16
Recognition Rate	87%

Greg Findlay, global market development at VoicePerfect, "We did the hard work to make deployment for our users easy."

There is no way around the "hard work" required to make dictation successful. But the work can be allocated; more customization on the back end means less of the user's time to make corrections or repeat recurring tasks. Assisting the user in making an educated choice as to the type and amount of customization required is essential for success.
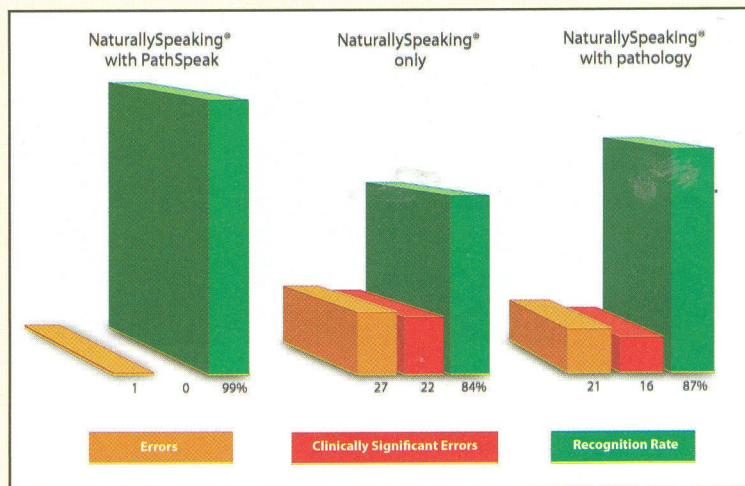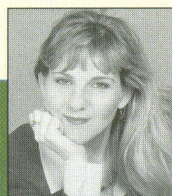
© 2005 Robin Springer

Figure 1: Total Error and Recognition Rate

Robin Springer is the president of Computer Talk (www.comptalk.com), a consulting firm specializing in the design and implementation of speech recognition and other hands-free technology services. She can be reached at (888) 999-9161 or info@comptalk.com.