

Dictation software for MDs improving but frustration still part of the program

Thomas Kovesi, MD

En bref

LE DR THOMAS KOVESI A RÉDIGÉ PRÈS DE 200 RAPPORTS DE CONSULTATION à l'aide d'un nouveau logiciel de dictée avant de préparer un aperçu critique de cette technologie pour le JAMC. Ces programmes s'améliorent, dit-il, mais armez-vous de patience : ils tapent incorrectement un mot sur cinq.

Computerized voice dictation sounds enormously attractive. After paying the bill for hardware and software, the cost of your dictations should virtually cease: computers don't need benefits such as unemployment insurance or a present at Christmas, and they don't take sick days.

But do these programs really live up to their promise? I've used voice dictation via the computer for virtually all of my consultant's reports for the past 4 months, and completed more than 180 reports so far. My experiences may be useful to other physicians trying to improve their own results with voice-dictation software.

The 2 key players in this market are NaturallySpeaking from Dragon Systems and ViaVoice from IBM. Both allow the use of continuous speech for dictation, rather than the . . . very . . . slow . . . discrete . . . speech required by earlier software. Both programs have roughly similar features and recognition accuracy.¹ To use them, start the program, place a headset microphone on your head and begin dictating. The results appear on the computer monitor.

Neither of these programs comes close to providing perfect speech recognition, so they employ strategies to improve their results. Before you use either program for the first time, you will have to read a chapter from a book to allow the program to learn your speech patterns. Each program gradually learns from its mistakes, so accuracy improves over time, and both learn to recognize words in context. This means that homonyms are usually spelled correctly — both programs should learn to spell "Jim may use Ventolin" and "John may return to gym class" correctly.

Physicians will also be able to train the computer to spell big words and/or jargon — ciliary dyskinesia is an example — correctly (see sidebar). Finally, each program transcribes only those words it knows how to spell. This can lead to some startling results. Don't be surprised if your program's first rendition of "cystic fibrosis" is something similar to "sixty-five roses."

Legal implications?

Both programs have some pronounced quirks. Neither will recognize anyone else's voice, and recognition accuracy plummets when you have a cold. With "training" the programs learn to recognize big words perfectly and most of the recognition errors will involve small words — "the" becomes "she," or "receive" becomes "received."

Physicians must be careful, because these errors could have major medicolegal implications. Consider the potential result if the sentence "We should definitely continue digoxin" becomes "We should definitely discontinue digoxin."²

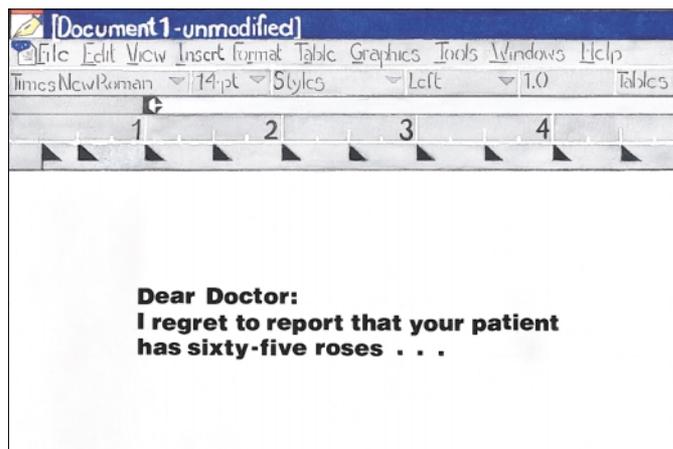


Features

Chroniques

Pediatric Respiriologist
Thomas Kovesi is an Assistant Professor of Pediatrics at the Children's Hospital of Eastern Ontario, Ottawa.

CMAJ 1998;158:1059-60





Since every word in your document will be spelled correctly, your word processor's spell-checker will be of minimal value; the grammar checker may not pick up errors either. This means every piece of dictation involving these systems require proofreading. I'm terrible at it, so my secretary still proofs the documents I dictate before they're printed. The current versions of these programs may be more useful for specialists who use the same terms in most of their letters.

Don't expect to run these programs on a 3-year-old Pentium. They require a fairly powerful computer — at a minimum a Pentium 133 for NaturallySpeaking and a Pentium 150 for ViaVoice. While both programs will run with 32 MB of RAM, I found that NaturallySpeaking's performance improved significantly with 64 MB of memory. Processor speed doesn't seem to be a critical factor — I haven't noticed much difference when using a faster computer with equivalent memory.

When the hardware is powerful enough, NaturallySpeaking is quick — it will happily record a phone conversation if I forget to turn the microphone off. However, I find dictation takes about 20% longer than when I use a dictaphone, partly because I find I can speak faster into a dictaphone than these programs seem capable of handling, and

partly because of the time taken to correct errors. NaturallySpeaking is a little faster than doing my own typing.

I usually dictate a sentence at a time, then correct the mistake(s) before going on. Many people prefer to dictate an entire document and then correct it afterwards. After four months, my system is reaching about 80% accuracy. This sounds impressive until you realize that this means 1 word in 5 — usually the small ones — will be incorrect. I don't know many users who achieve more than 90% accuracy, despite the programs' claims.

Placement of your headset microphone is a critical factor if you want to achieve satisfactory performance. It is supposed to be placed just below and in front of your mouth, with the "front" of the microphone facing your mouth. A minor adjustment in microphone placement may do more for accuracy than spending several hundred dollars on new hardware.

The most crucial principle to remember if you wish to achieve satisfactory results is that the less you dictate, the fewer errors the software will make. I find that using NaturallySpeaking has given me a great deal of "positive reinforcement" that encourages brief reports. I use pre-prepared skeleton documents (templates), and fill in the blanks. This saves time: unlike conventional dictation, you don't have to say the standard words used in every document, such as "Dear Dr." It can also give consultants' letters a consistent, easy-to-read look that is difficult to achieve if you use several transcriptionists.

NaturallySpeaking lets users move around a document using voice commands: "Move to the third word on the fourth line." I find it faster to navigate with the mouse, particularly because the program sometimes misunderstands the voice command. IBM's ViaVoice lets users create "voice macros" that type a sentence or more once you say a keyword.¹

Should you switch to one of these programs for your dictation? If you're willing to spend 20% longer on the work in return for a significant reduction in overhead, these programs are certainly worth trying. Dictating will be more arduous than with the dictaphone, since you will keep glancing at the screen and making corrections, but my reports have become shorter and none of my colleagues has complained about that. They are also better organized and at least as informative as before. And if I do feel like being more expansive in a particular letter, neither my PC nor my budget will care!

Dear Dr.: This letter was dictated . . .

Dr. Thomas Kovesi dictated the following paragraph using NaturallySpeaking software by Dragon Systems, version 1.0. This is how it appeared on his screen.

"Neither of these programs come anywhere close to perfect speech recognition, so they using number of strategies to improve their results. Before you use either program for the first time, you must alleviate chapter from a book to your computer, to let the program learned your speech patterns. Each program gradually learns from its mistakes, so accuracy improves over time. Each program also learns to recognize words in context, so, names are usually spelled correctly-the program should learn to spell "Jim me use Ventolin" and "John may return to Jim class" correctly. You can train the computer to spell big words and/or jargon (like "ciliary dyskinesia") correctly. Finally, each program trance drives only words it knows how to spell. This can lead to some startling results. Don't be surprised if your programs first rendition of 'cystic fibrosis' is something similar to 'sixty-five roses.' "

References

1. Gerding D. Speak Your Mind. *PC Computing* 1997;10(11):292-304.
2. Woram J. PC: Talk To Me! *Windows Magazine* 1997;8(11):208-20.