Voice Recognition Software: Comparison and Recommendations

Use of voice recognition software is under consideration by medical office administrators nationally. Administrators have long searched for alternatives to the expense, error rate, and record-completion delays associated with conventional transcription. It is no wonder that, with the recent advances in voice recognition software, medical transcriptionists are looking at this emerging technology as a powerful way of accomplishing essential record-keeping tasks.

This report investigates four of the leading voice recognition applications to determine whether this technology has become a practical option and to determine which application is the best choice. And so that this report and further study of the software can be better understood, an introduction to the subject of voice recognition software follows.

Introduction to Voice Recognition Technology

Several different voice recognition products currently exist in the marketplace, and viable choices are greater in number than they were only a few years ago. Rapid changes have been fueled by the ever-increasing power and plummeting prices of desktop systems. Though room for improvement still exists, accuracy has advanced tremendously in a stunningly short time.

Brief history

The first software-only dictation product for PC's, Dragon Systems' DragonDictate for Windows 1.0, using discrete speech recognition technology, was released in 1994. Discrete speech is a slow, unnatural means of dictation, requiring a pause after each and every word [11]. Two years later, IBM introduced the first continuous speech recognition software, its MedSpeak/Radiology. These systems often had five-figure price tags and required very expensive PCs. Continuous speech technology allows its users to speak naturally and conversationally, relieving much of the tedium of discrete speech dictation [11].

Dragon Systems made an enormous stride in June, 1997, when it released NaturallySpeaking, the first general-purpose continuous speech software program. Much more affordable than earlier programs, it brought the realm of continuous speech recognition to a much wider range of users. Two months later, IBM released its competing continuous speech software, ViaVoice [10].

Stringent demands

Much is demanded of speech recognition programs. Accuracy is critical, and speed is essential to any effective program. Added to these challenges are the enormous variance that exists among individual human speech patterns, pitch, rate, and inflection. These variations are an extraordinary test of the flexibility of any program. Voice recognition follows these steps:

1. Spoken words enter a microphone.
2. Audio is processed by the computer's sound card.
3. The software discriminates between lower-frequency vowels and higher-frequency consonants and compares the results with phonemes, the smallest building blocks of speech. The software then compares results to groups of phonemes, and then to actual words, determining the most likely match.
4. Contextual information is simultaneously processed in order to more accurately predict words that are most likely to be used next, such as the correct choice out of a selection of homonyms such as merry, marry, and Mary.
5. Selected words are arranged in the most probable sentence combinations.
6. The sentence is transferred to a word processing application [11].
Power devourers

With all of the complex selections and tremendous flexibility demanded of voice recognition software, it is small wonder that considerable computer muscle is required to run these programs. To take fullest advantage of current speech recognition programs, a PC with a minimum of a 300 MHz Pentium II processor is recommended. A separate 16-bit SoundBlaster-compatible card is also advisable, because the sound cards that are bundled as part of a PC's motherboard can produce inferior results with voice recognition software [4].

Realistic reminders

The technology has advanced impressively over the last year, with programs variously offering smarter speech recognition engines, larger active vocabularies, integration with the most popular word-processing programs, and improved accuracy. This report sorts through these to find the most accurate program and the best value available, and determines if the accuracy is acceptable at this time [4]. It is essential to remember the following:

• While voice recognition software has made enormous strides, it is not perfect. Dictated records, particularly in the first few weeks of use, must be sufficiently proofed while onscreen.
• Since medical and legal requirements for record keeping are exacting and extensive, considerable dictation is required. Dictation using voice recognition software is like many other things: practice makes all the difference. Tests by PC Magazine Labs showed that increased experience with dictation and the software clearly increased accuracy [3]. Be prepared to invest a few weeks of dictation time and practice with the software in order to see enhanced accuracy.

Requirements for the Purchase of Voice Recognition Software

Based upon stated preferences and system specifications, the following conditions have been established:

• Continuous speech recognition software must provide is preferred, rather than the slower, more unnatural and lower-priced discrete speech recognition software also on the market.
• The application must run on a Pentium-powered PC under Windows 95, and be capable of integration with Microsoft Word97.
• The software program must be easily and successfully installed by any intermediate-level computer user in the office.
• The program must be one that can be learned and customized reasonably quickly by nearly anyone in the office.
• The cost limit is $1,500.

Points of Comparison

The different voice recognition software programs compared are Dragon Systems' NaturallySpeaking 3.0 Preferred Edition, IBM ViaVoice 98 Executive, L&H Voice Xpress Plus, and Philips FreeSpeech98. Discussion of Dragon Systems' NaturallySpeaking will also include its Medical Suite.

Eight categories of comparison will be made in order to effectively evaluate these competing programs: (1) accuracy; (2) minimum system requirements; (3) capacity to manage a specialized medical vocabulary and medical records; (4) integration with Microsoft Word; (5) ease and speed of installation, customization and use; (6) industry ratings and awards; (7) inclusion of microphones, and (8) cost.

Accuracy

Accuracy is the single most significant consideration; without it, the program is useless. Dragon Systems' NaturallySpeaking 3.0 scored highest on all of the accuracy tests performed by PC Magazine and was unequivocally selected as the Editors' Choice. In their tests, the average accuracy was 91% and at times was considerably greater [1].
Average accuracy for L&H Voice Xpress was 87% [2]. Accuracy for IBM's ViaVoice tested at 85% [14], and Philips FreeSpeech98 was 80% [15].

At first glance, these percentages, particularly the top two, may not seem significantly different. Consider, however, that for every 1,000 words, an accuracy rate of 87% means that 130 words must be corrected. An accuracy rate of 91% represents an average of 90 errors per 1,000 words, while an 80% rate means that 200 out of every 1,000 words must be corrected.

Thousands of words are dictated daily in this practice. Time is scarce and precious. Medicolegal conditions mandate that records must be exhaustively thorough and accurate. Under these rigorous circumstances, with every percentage point counting heavily, Dragon Systems' NaturallySpeaking yields the highest accuracy.

**Minimum system requirements**

All four programs run on Pentium-powered PC's utilizing Windows 95, 98 or NT 4.0 and require 16-bit SoundBlaster-compatible sound cards. Random access memory (RAM) requirements for software run under Windows NT are higher for all of these programs [5].

- Dragon Systems' NaturallySpeaking requires a Pentium/133MHz processor or higher, 32MB of RAM, and 180MB of hard disk space [5].
- IBM ViaVoice 98 requires a Pentium/166MHz with MMX (multimedia chip) or higher, 32MB of RAM, 180MB of hard disk space, and 256K L2 cache [5].
- L&H Voice Xpress Plus requires a Pentium/166MHz with MMX, 40MB of RAM, and 130 MB of hard disk space [5].
- Philips FreeSpeech98 requires a Pentium/166MHz processor, 32MB of RAM, and 150MB of hard disk space [5].

<table>
<thead>
<tr>
<th>Software</th>
<th>CPU</th>
<th>RAM</th>
<th>Hard Disk Space</th>
<th>L2 Cache</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dragon</td>
<td>Pentium/133 MHz</td>
<td>32  MB</td>
<td>180 MB</td>
<td>none</td>
</tr>
<tr>
<td>IBM ViaVoice</td>
<td>Pentium/166 MHz-MMX</td>
<td>32  MB</td>
<td>180 MB</td>
<td>256 KB</td>
</tr>
<tr>
<td>L&amp;H</td>
<td>Pentium/166 MHz-MMX</td>
<td>40MB</td>
<td>130MB</td>
<td>none</td>
</tr>
<tr>
<td>Philips</td>
<td>Pentium/166 MHz</td>
<td>32  MB</td>
<td>150 MB</td>
<td>none</td>
</tr>
</tbody>
</table>

It is important to recall that, as noted earlier, significantly greater system resources are recommended to optimize performance. Given the sufficient system resources, none of these software programs should present a problem for the existing system.

**Capacity to manage a customizable, specialized medical vocabulary**

Medicine in general, and each medical specialty in particular, have their own complex, specialized vocabularies.

- Dragon Systems NaturallySpeaking offers a so-called Medical Suite targeted to medical professionals and specified as an alternative to transcription. Marketing materials state that an extensive vocabulary of thousands of words, including medical procedures, terms, drugs, diagnoses and symptoms, are included. The software allows creation of multiple vocabularies for specialty customization if desired [8].
IBM offers add-on VoiceType Vocabularies for use with ViaVoice. The medical vocabularies available are for Emergency Medicine Dictation and Radiology Dictation. No other specialty customization is available [13].

L&H Voice Xpress and Philips FreeSpeech98 do not offer medical vocabularies, either as add-ons or bundled with the software [9, 12].

Two of the four companies offer a product that provides medical terminology. IBM's emergency room and radiology add-on software is not applicable to the dictation needs of obstetric and gynecologic practices, for example. Dragon Systems' NaturallySpeaking Medical Suite offers the same voice recognition technology as the previously mentioned NaturallySpeaking Preferred Edition, with the addition of extensive customizable medical terminology that can be tailored to other specialty practices.

Integration with Microsoft Word

All four programs integrate with Word97 and can therefore be used with existing word processing software [5].

Ease and speed of installation, customization and use

Each of the four programs uses "wizards" to install and configure hardware, and all programs support macros for frequently used phrases.

- Dragon Systems' NaturallySpeaking uses its wizard to train the system to recognize the user's voice within 4 minutes. Material is provided so that about 30 minutes of reading aloud will improve accuracy [5]. Electronic medical documents can be analyzed automatically to "learn" new specialized terms and proper names. Its CommandWizard feature enables any user to create medical-specialty macros. Commonly used and required medical forms, electronically stored, can be readily called up and the user is prompted to fill out each section of a form [8].
- IBM's ViaVoice also trains the system by means of reading from selected texts for about 30 minutes, and its wizard adjusts microphone and speaker volume levels [5].
- L&H Voice Xpress Plus directs the user to read chapters of a book, and in PC Magazine's tests, about 75 minutes was required for the process [5].
- Philips FreeSpeech98 directs the user to read selected text for about 15 minutes; ten training topics are available for the user's review [15].

Installation of all of the programs appears straightforward, and the initial basic "training" is not excessively time-consuming for any of the products. While all provide macros, the medical customization features of Dragon Systems' product are considerably greater. Though they will initially require more time and document input, accuracy is increased, and for this reason, Dragon's software is recommended in this comparison.

Industry ratings and awards

Only one of these products refers to and lists awards on its web site, and that is Dragon Systems' NaturallySpeaking. None of the other three products has any such mention anywhere on its site, nor do any awards or industry recognition show up on multiple web searches for the products.

Dragon Systems' web site lists over fifty awards, some of which are listed here:

- *PC Magazine*, Editors' Choice, October 1998; this particular article is referenced several times in this report [1, 7].
- *PC/Computing*, Time Capsule - The 12 Best PC Products on the Planet: Input Device Category - August 1998 [7].
- *BYTE* Best - January 1998 [7].
- *BusinessWeek*, The Best New Products/Software - January 1998 [7].
- *PC/Computing*, 5 Star Rating - November 1997 [7].
While industry recognition and journalistic evaluations are not the only considerations, Dragon Systems boasts an impressive list of awards and ratings by prestigious periodicals.

Inclusion of microphones

As previously noted, a microphone is necessary for capture of spoken words.

- Dragon Systems ships with a VXI Parrott 10-3 microphone; *PC Magazine* notes that it is comfortable and performs well [5].
- IBM's ViaVoice and L&H Voice Xpress Plus both provide an Andrea NC-80 microphone, which *PC Magazine* states is not as comfortable as the XVI Parrott 10-3 [5].
- Phillips FreeSpeech98 does not include a microphone; it recommends its own SpeechMike at an extra cost of $69.95 [5].

None of these is a make-or-break detail, but Dragon Systems has a slight edge with the reviews provided by *PC Magazine*.

Cost

Highly significant price differences exist among these programs.

- The Dragon Systems' NaturallySpeaking Preferred Edition tested by *PC Magazine*, October 1998, retails for $179 when purchased directly from Dragon or through resellers. Rather than purchasing this edition for a medical practice, NaturallySpeaking Medical Suite is available for $995. An Add-On Medical Specialty Vocabulary is $49. One year of 800-number telephone support for all products is an additional $199, for a total cost of $1,243, exclusive of tax and shipping costs, for the Medical Suite [6].
  - IBM's ViaVoice 98 Executive software program costs $150, and the medical specialty add-ons are $240. However, these add-ons are for emergency medicine and radiology only [13].
  - L&H Voice Xpress Plus is $70 [5].
  - Philips Free Speech98 costs $39 and includes no microphone. A Philips SpeechMike can be ordered for $69.95, for a total cost of $108.95, exclusive of tax and shipping costs [5].

L&H offers the best price by far. IBM and Philips are roughly in the same ballpark. Dragon Systems' Preferred Edition is more expensive at $200, but not significantly so. The only customizable medical software program is Dragon Systems' Medical Suite, which, at $1,243, is over ten times the cost of Philips' software, though it includes one year of technical support.

Summary

From business, medical, and legal perspectives, the creation and maintenance of accurate, complete records are crucial. The primary downside to such thorough record-keeping includes: (1) the time required for dictation, (2) the costs in and inherent hassles of finding and hiring a competent medical transcriptionist, (3) the necessary delays between dictation and actual availability of the transcribed records, and (4) the time needed to proof and correct the transcriptionist's output.

To date, the weakest link in speech recognition technology has been accuracy. This is fast changing, and current software programs have significantly improved within the last year. Can a voice recognition software program eliminate some of the problems occurring in conventional medical transcription? The following conclusions will help answer this question in the recommendation that follows:

1. All of the programs specify system requirements that are well within the parameters of the existing system.
2. All of the programs integrate with the existing word processing software, Microsoft Word97.
3. All of the programs can reasonably be installed by the average user.
4. Dragon Systems NaturallySpeaking Medical Suite is by far the most expensive voice recognition program. While it is $1,243, including one year of technical support, the other three programs are all under $200, exclusive of support.
5. Philips does not include a microphone with its software as do the other three software companies, but purchase of one does not increase the total cost appreciably. Dragon Systems' microphone is considered more comfortable than the other microphones tested by *PC Magazine*.

6. Dragon Systems' NaturallySpeaking has accumulated a lengthy list of awards; no awards were found for the other three programs.

7. Dragon Systems' NaturallySpeaking Medical Suite with Add-On Vocabularies is easily customizable to specific needs of different practices for specialized medical vocabulary and medical forms.

8. Dragon Systems' NaturallySpeaking technology is the most accurate of the four programs tested.

9. Although Dragon Systems' NaturallySpeaking is the most expensive, it offers the best function while the other options considered are barely adequate.

10. The best choice of the four applications considered is Dragon Systems' NaturallySpeaking.

**Recommendation**

Dragon Systems NaturallySpeaking Medical Suite is strongly recommended for its superior accuracy, powerful customization features, and industry recognition and awards. No other product comes close, and its strong advantages justify its higher price. Once the program has been customized, and the user has dictated for several weeks and become familiar with the software, acceptably accurate transcription and instantly available medical records should be possible with NaturallySpeaking Medical Suite, solving some of the record-keeping problems faced by this medical practice.

**Literature Cited**

All references are found online:


http://www.io.com/~hcexres/textbook/