

11-2000

Voice Recognition Software

Bobby E. Waldrup


University of North Florida, bwaldrup@unf.edu

Jeffrey E. Michelman

University of North Florida, jeff.michelman@unf.edu

Maxine L. Person

Follow this and additional works at: https://digitalcommons.unf.edu/bacc_facpub

 Part of the [Accounting Commons](#), [Computer Sciences Commons](#), and the [Finance and Financial Management Commons](#)

Recommended Citation

Waldrup, Bobby E.; Michelman, Jeffrey E.; and Person, Maxine L., "Voice Recognition Software" (2000).
Accounting and Finance Faculty Publications. 7.
https://digitalcommons.unf.edu/bacc_facpub/7

This Article is brought to you for free and open access by the Department of Accounting and Finance at UNF Digital Commons. It has been accepted for inclusion in Accounting and Finance Faculty Publications by an authorized administrator of UNF Digital Commons. For more information, please contact [Digital Projects](#).
© 11-2000 All Rights Reserved

What if you could talk to your computer?



In Brief

A Promising Technology Not Ready for Prime Time

Voice recognition software (VRS) allows computer users to control applications with voice commands instead of the traditional keyboard and mouse. The technology is being heralded as the next generation of human-computer interface, promising to make computer applications more efficient and user-friendly.

The authors introduce the three major off-the-shelf VRS packages, discuss possible uses for CPAs, and chronicle their difficulties in using the technology and adapting to its requirements and shortcomings. Those hoping for immediate increased efficiency in all aspects of usage will likely be disappointed; VRS delivered on certain promises but fell far short of the authors' expectations in others. Prospective users are warned that off-the-shelf packages are not necessarily ready to fulfill all the needs of the demanding CPA.

Voice Recognition Software

By Bobby E. Waldrup, Jeffrey Michelman, and Maxine L. Person

Now that Y2K has come and gone, CPAs can consider investing in new computer technologies, searching for software that offers innovative new capabilities. Voice recognition software (VRS) is a relatively new technology with a great deal of promise. VRS allows professionals to control computer applications with voice commands in lieu of the traditional keyboard and mouse. The ultimate promise of the technology is to allow humans to interact with computers more naturally and efficiently; however, the future is a long way off. Current VRS offerings provide significant benefits, but not without accompanying limitations and drawbacks, especially for technical professionals.

Capabilities

VRS provides for the two broad functions of voice-activated dictation and macro-level control over applications. Dictation is the most commonly used function. With a headset attached to the computer or a digital voice data-collector similar to the hand-held tape recorder, users can directly record unlimited dictation into popular word-processing programs such as Microsoft Word and Corel WordPerfect.

In the second most commonly used VRS function, voice commands—rather than a mouse or keyboard—are used to operate applications. The user can also voice-activate applications to save, print, and move files and perform mathematical functions in spreadsheet and database applications.

Vendors. The majority of the VRS market share is controlled by three vendors: IBM (ViaVoice), Lernout & Hauspie (Voice Xpress), and Dragon Systems (Dragon Naturally Speaking). [At the time of writing, Lernout & Hauspie had acquired Dragon Systems as a wholly owned subsidiary. Although L&H Senior Product Director Bill DeStaefanis pub-

licly announced that both brands will continue to be sold, prospective adopters should consider the effect the merger may have on future product offerings.

The three vendors make similar claims about the capabilities of their respective brands. All indicate spelling and grammar accuracy upwards of 95% for dictation applications. They each state that initial voice recognition will require 12-15 minutes training time but that longer use will allow for greater accuracy. According to vendor marketing claims, the software should be capable of integration with other Windows-based (and, in the case of the IBM brand, Macintosh-based) business applications.

The purpose of this article is to share the authors' findings with respect to these vendor claims. We argue that accountants should use caution before adopting this new technology for intensive professional use.

Current Trends in Practice

All areas of accounting practice have been, to some degree, revolutionized by automation. This can best be illustrated by the explosion of prepackaged PC-based software that can assist in tax preparation and general ledger, database, spreadsheet, and work-paper delivery systems. These innovations are in addition to commonly used office-based software for word processing, e-mail, and presentation. Accounting applications, while producing significant efficiencies in a professional office, still must generally be operated in a stand-alone fashion. The efficiencies gained by computer technologies can be enhanced by the user's ability to move data fluidly among different packages.

Practitioners may be tempted to adopt VRS as a means of efficiently toggling between various accounting and office applications. VRS vendor marketing efforts portray office

EXHIBIT 1 VRS VENDOR CLAIMS

Vendor Claims *	Lernout & Hauspie Voice Xpress v.4	Dragon Naturally Speaking v.4	IBM ViaVoice v.7
Ease of installation and learning curve	Learns user's voice in under 10 minutes	Dictating capable within 15 minutes	Easy to train, create personal voice model
Ease of use	Say it your way, accurately turns user's voice to text	True continuous speech, just speak to your computer naturally	Simple to use, use natural voice and watch words appear
Software performance: Dictation functions	Up to 140 words or more per minute	Achieve greater than 95% accuracy, 230,000 word total vocabulary	Dictate using natural language technology into Word, Excel, and Outlook
Software performance: Command functions	Voice-enable hundreds of other applications	Editing, formatting, printing, and copying into Word and WordPerfect	Create voice navigation macros using Navigation Macro Creator

* As adapted from vendor verbiage

Editor's Note: All the vendors have recently released updated versions of the software. The authors did not test the claims of these upgrades.

professionals moving effortlessly among a wide variety of these applications using natural language technology (NLT) without resorting to keyboard or mouse commands.

The three dominant VRS packages mentioned above were tested to determine their applicability to various accounting related tasks. These packages share the following commonly advertised features:

- Dictation training in less than 45 minutes
- Dictation directly into popular word processing programs
- Natural language technology for the Microsoft Office suite
- Remote transcription capability
- Limited command features over some Windows applications.

Analysis

To facilitate analysis, each of these software packages was installed on a Windows desktop PC. The computer's hardware system significantly exceeded the minimum requirements necessary according to all vendor specifications. The investigators analyzed the performance of the VRS pack-

ages as a group based upon the following criteria (See *Exhibit 1*):

- Ease of installation
- Learning curve
- Ease of use
- Software performance for both dictation and command functions.

Each of the packages was evaluated on an A-F scale by two separate evaluators. These grades were then averaged to produce an overall technology

grade, rather than a comparative ranking of the individual software packages, which generally performed similarly (See *Exhibit 2*). The grade represents an evaluation of the software's performance as it applies to the typical CPA. This grading scale is inherently subjective and is meant only to serve as broad guidance (See *Sidebar*). Recommendations of specific software packages are not implicit in the evaluations.

Ease of installation, learning curve, ease of use. The investigation found that each of the packages installs quickly, and learning to get around the programs is easy and relatively painless. Training protocols are simple, and a PC-literate professional will have any of the packages running within an hour. One concern is that the programs had difficulty learning specialized vocabulary, such as proper nouns and specific accounting terms.

Software performance: dictation and command functions. VR technology excels at quickly and accurately recognizing common speech patterns.

(Continues on page 49)

Specializing the vocabulary of these VRS packages is not as easy as advertised.

EXHIBIT 2
VRS TECHNOLOGY REPORT CARD

Feature	Overall Grade	Notes
Ease of Installation		
Ease of software installation	A+	Typical minimum system requirements: Pentium II, Windows 95/98, 48 MB RAM, 200 MB free disk space.
Learning Curve		
Quality of user guides	A-	Adequate user guides, but must understand the program language.
Training time before first dictation	A-	Generally short, but varied by speaker and environment.
Training time for special vocabulary	D	Frustratingly inaccurate with proper nouns and complex vocabulary.
Ease of Use		
Headset comfort	C+	Users never adjusted to the feeling of being tethered to the computer and required frequent adjustment.
Intuitive for user interface	C+	Make no mistake, the user must learn another language.
Software Performance		
Ability to compensate for background noise	C-	User must be in a library-like setting for optimal use.
Ability to increase productivity	F	On balance, this technology reduces productivity.
Dictation Functions		
Dictation capabilities into other Windows applications	D	Claims of seamless integration with other non-word-processing applications are greatly exaggerated.
Dictation accuracy: common	B	As a rule, sentences filled with five-letter words are a strength of these packages.
Dictation accuracy: specialized	F	Packages never learned many proper nouns despite repeated training attempts and were weak in learning number dictation.
Dictation speed	A+	All of the VRS packages could type as fast as the user could speak.
Command Functions		
Editing functions	D	The packages falter with these functions. Certain common commands were consistently misinterpreted.
Ability to move among applications	F	The ability to toggle into, out of, and among applications is disappointing.
Overall Evaluation	C/C-	Most CPAs would hardly choose to hire staff with this level of performance; therefore, it is unlikely that they will choose to adopt technology at this level of performance.

EXHIBIT 3 FUTURE VRS APPLICATIONS

Position	Organization	Potential Applications
Tax attorney (non-CPA)	National law firm	<ul style="list-style-type: none"> ■ Improved efficiency through enhanced quality dictation ■ Integrate into decision tree process of document assembly
Tax consultant	Big Five firm	<ul style="list-style-type: none"> ■ Improved efficiency through dictation of tax terminology ■ Seamless integration with Microsoft Office and Lotus Notes
Tax manager	Large local firm	<ul style="list-style-type: none"> ■ Improved quality of dictation ■ Integration with intelligent agents to complete tax research
Owner (2)	Sole practitioner	<ul style="list-style-type: none"> ■ Improved efficiency through enhanced quality dictation ■ Navigation aid to entering data into schedules ■ As an aid to entering search criteria in tax research
Audit senior manager (2)	Big Five firm	<ul style="list-style-type: none"> ■ Entry of data into a personal digital assistant (PDA) ■ Documentation of client interaction ■ Foreign client work-paper translation ■ As an aid in computer-based training or learning ■ Conversion of text to voice in order to improve comprehension
Audit partner	Large local firm	<ul style="list-style-type: none"> ■ Improved efficiency through enhanced application integration ■ Must have 100% data integrity
Finance manager	Transportation company	<ul style="list-style-type: none"> ■ Reduce repetitive stress injuries ■ Reduce barriers to knowledge management

(Continued from page 44)

Each VRS package touts dictation accuracy upwards of 95% using commonly recognized vocabulary. However, once the user deviates from even the most basic of vocabulary, the technology begins to falter. Concerns have to do with three basic failures: Pervasive background interference during use, Dictation into non-word-processing applications, and Dictation of specialized vocabulary.

Ambient background noise during dictation significantly reduces the accuracy of each of the software packages. Heating and cooling fans sometimes cause significant error rates in dictation. If a telephone rings or an intercom buzzes, each of the software packages will continue to respond, typing gibberish. In addition, if the user is hoarse or congested, the accuracy of dictation can be severely hampered.

Each of the packages allows for dictation into proprietary notepad programs, which can be translated into any of the major word-processing pack-

ages. Each also allows for dictation directly into other application programs, such as databases, e-mail, and spreadsheets. We found this dictation application to falter significantly with regard to accuracy and speed. For example, one investigator with a southern accent found that dictating the words "effort" and "forth" into an Excel spreadsheet would repeatedly cause the cursor to jump to the cell "F4." Another investigator had difficulty with the command functions; when commanded to "delete that," the software consistently typed "believe that."

The investigators also found that specializing the vocabulary of these VRS packages is not as easy as advertised. Dictation of special terminology such as SAS, FASB, and client corporate and individual names quickly becomes burdensome. Several proper names were never correctly learned by the packages, despite concentrated and repeated specialized training that is far beyond the patience and available time of most CPAs. While each of the appli-

cations allows for numerical and dollar-denominated nouns (e.g., \$1,232.45), vendor-specific pronunciations of all numerical amounts must be used for the software to comprehend.

The promise of VR technology to create a voice-controlled command environment in Windows-based applications is disappointing. Using voice commands in lieu of the mouse is an arduous process. Often, out of sheer aggravation, the investigators would simply reach for the mouse to complete a simple command such as printing. While it does appear that the packages can substitute voice commands for practically all mouse and keyboard functions, in practice this feature would only appear to be of benefit to individuals that could not use or encountered extreme physical difficulty when operating a traditional mouse and keyboard.

Overall Evaluation

The VRS report card (Exhibit 2) indicates a wide variability of performance in different functional areas. Although

there was slight variation in performance among the three products, this did not impact the overall assessment. Perhaps the most surprising result of this study is that the overall evaluation of VRS software for CPAs' tasks is between a C and C-. Most CPAs would hardly choose to hire staff with this type of performance record; therefore, it is unlikely that they will choose to adopt technology at this level of performance.

Current Practice Needs Using New Technology

As currently packaged, the only real strength of VRS technology lies in simple dictation. This is unfortunate for CPAs, because professional usage would not generally focus upon these basic tasks. Currently, the needs of CPAs revolve around e-mail applications, control of quantitative office software (i.e., spreadsheets and databases), client-centered work-paper production, and control of various proprietary professional accounting packages. Even allowing an optimistic 95% accuracy rate in all applications would leave open the possibility of serious error.

Proprietary accounting-specific VRS is both possible and probable in the near future. Specialized software can be written effectively for profession-specific applications, as demonstrated in the numerous packages available for health care tasks such as medical transcriptions, patient billing records, and pharmaceutical prescriptions. However, this kind of specialized software is not currently available for the accounting profession; off-the-shelf applications such as those examined above are the only available alternative for CPAs hoping to adopt VR technology.

Room for Improvement

In order to examine accountant needs for high-quality VRS, the authors interviewed eight practicing CPAs and a tax attorney. All nine had some knowledge of the VRS but none had ever tried it. The consensus was that

Each VRS package touts dictation accuracy upwards of 95% using commonly recognized vocabulary.

anything which could improve the efficiency of various aspects of their work would be welcome (The results are presented in *Exhibit 3*). Perhaps the most pervasive issue for accountants, their employees, and their clients is the possibility of suffering from repetitive stress injuries, such as carpal tunnel syndrome. Hence any technology that could reduce the incidence of these debilitating injuries would be welcomed.

The group generally agreed that the keyboard and mouse have become barriers to data entry through either lack of typing skills, limitations of time and place, or the inability to foster an environment for the dialogue necessary for Internet or tax research. For these reasons, the group was excited about the possibilities of VRS—if its accuracy were improved, its integration with applications were enhanced, and it could facilitate better research capabilities through knowledge management. Finally, several interviewees noted that VRS needed to be used with cau-

tion at client sites due to the confidential nature of the data being entered.

The Future

VRS is a relatively new technology. Off-the-shelf packages can be initialized, by a technologically savvy user, in under an hour and be trained, with a slight learning curve, to handle verbal dictation with an accuracy upwards of 95%. However, CPAs need significantly more complex forms of automation. Currently available packages will not meet the expectations of the typical accountant. A certain level of professional skepticism is warranted; the authors recommend that CPAs delay adoption of VRS until it is further developed for non-word-processing applications by the major vendors or until tested, accounting-specific versions of the software are available.

In looking toward the future, we feel that it is particularly important to put the state of VR technology into perspective. During the early 1980s it would have been hard to find a significant number of CPAs who would champion the value of electronic spreadsheets or, for that matter, PC technology in general. Furthermore, many of us welcomed the AICPA partnership with the Grammatik software in the early 1990s, which made grammar checking much easier. We clearly envision the development of an accounting-based VRS dictionary that would provide improved functionality similar to those applications currently available to the medical and legal professions.

The long-term future development of VR technology hinges upon developing a neural network technology that mimics the way the brain interprets speech. Neural networks use whole word analysis rather than the current approach, which segments speech patterns into smaller chunks. We believe that once computer science advances to a more brain-like interpretation matrix, VR technology will indeed become the new standard of data entry. Thus, even though we recommend against early

RESOURCES

The scope of this article has been to evaluate the overall suitability of VRS for accounting tasks. The authors have not attempted to compare and contrast the individual personalities of the three major brands available on the market today. Before adopting a VRS solution, users should research the specific requirements and capabilities of each package to see if one is particularly suited to the task or work environment.

Vendor Resources:

Dragon Systems Inc.: www.dragonsystems.com; (800) 437-2466

IBM ViaVoice: www.ibm.com/viavoice; (800) 825-5263

Lernout & Hauspie: www.lhsl.com; (800) 380-1234

Websites for many less prominent vendors can be found at dir.yahoo.com/Business_and_Economy/Shopping_and_Services/Computers/Software/Voice_Recognition/

Independent Resources:

PC Magazine (Nov 5, 1999) presented an in-depth comparison of these three VRS products, as well as references to numerous individual reviews for each brand. Expanded and updated reviews can be found online at www.zdnet.com/eweb/ (under the "Reviews" option, search for "voice recognition"). □

adoption of VR technology, we strongly recommend that the accounting profession continue to follow this emerging standard.

CPAs might consider the day when VR technology facilitates scripted client interviews, whereby clients can easily prepare auditor checklists that are non-intrusive, yet accurate. Great strides in VR technology have already been achieved in other professional disciplines. Although they have yet to achieve the productivity and accuracy rates that CPAs require, advances in the technology will have a profound effect on the profession. □

Bobby E. Waldrup, PhD, CPA, is an assistant professor of accounting, **Jeffrey Michelman, PhD, CMA, CPA**, is associate dean of accounting, and **Maxine L. Person** is an MBA candidate, all at the University of North Florida, Jacksonville.
