Combining Speech Recognition Software With Digital Imaging and Communications in Medicine (DICOM) Workstation Software on a Microsoft Windows Platform

Randy Ernst, Walter Carpenter, William Torres, and Scott Wheeler

This presentation describes our experience in combining speech recognition software, clinical review software, and other software products on a single computer. Different processor speeds, random access memory (RAM), and computer costs were evaluated. We found that combining continuous speech recognition software with Digital Imaging and Communications in Medicine (DICOM) workstation software on the same platform is feasible and can lead to substantial savings of hardware cost. This combination optimizes use of limited workspace and can improve radiology workflow.

THE INTRODUCTION of computers into the radiology workplace can lead to greater productivity. However, implementation without careful planning may lead to decreased study turnaround and reduced acceptability by radiologists. Design of diagnostic workstations and radiology reading rooms, with particular attention given to lighting conditions, noise reduction, and optimal use of limited workspace, can greatly improve and facilitate radiology workflow.[1] Continuous speech recognition has been shown to be cost-effective and to markedly improve report turnaround.[2,3] Software technology has now improved to the point that the radiologist can speak at a natural, faster pace, thus making these systems more feasible.[2-5] By linking computer speech recognition to digital radiography, the radiology information system (RIS), and the picture archiving and communication system (PACS), the Mayo Clinic in Scottsdale was able to dramatically reduce radiology turnaround. In their case, the interval beginning at the moment that x-rays pass through the patient's chest to the moment that images are finalized and the radiologist's report is available anywhere in the clinic or hospital can be as little as 10 minutes.[6]

In the spring of 1999, Power Scribe (Lernout & Hauspie Speech Products, Burlington, MA) continuous speech recognition transcription software was implemented at the Emory Clinic and Hospital Radiology Departments. Since that time its use has been expanded to other Emory outpatient centers and to Grady Memorial Hospital. RIS, DECRad (IDX Corp, Burlington, VT), hospital information system (HIS), and PowerChart (Cerner Corp, North Kansas City, MO) query programs were placed on workstations with the speech recognition software. The network was also upgraded to switched 100BaseT. More than 30 different workstations are in use by residents and faculty for clinical dictations. Migration to digital review of ultrasound, computed tomography (CT), and magnetic resonance (MR) images is in progress. Clinical review workstations are installed, resulting in decreased work areas with additional monitors and computers generating more heat. In many cases the clinical review workstations are widely separated from the speech recognition computer. This lack of availability may hamper workflow and reduce acceptability by radiologists.

Our objective was to reduce the number of computers in our reading room by combining commercially available continuous speech recognition software with Digital Imaging and Communications in Medicine (DICOM) workstation software on the same platform. Optimal computer hardware, including processor speed and system random access memory (RAM), was evaluated. This project demonstrates advantages, disadvantages, and costs of combining computer hardware versus maintaining separate computers for each application.

From the Department of Radiology, Emory University Hospital, Atlanta, GA; and the University of Texas Medical Branch at Galveston, TX.

Supported in part by Lernout & Hauspie Speech Products, Burlington, MA.

Address reprint requests to Randy D. Ernst, MD, Department of Radiology, Emory University Hospital, Room A-141, 1364 Clifton Rd, NE, Atlanta, GA 30322. E-mail: rernst@massmed.org.

0897-1889/01/1402-1048$35.00/0
doi:10.1053/jdim.2001.23859

MATERIALS AND METHODS

Windows NT/Windows 2000 (Microsoft, Redmond, WA) were selected as the operating systems. Power Scribe was the selected continuous speech recognition software and eFilm (University Health Network, Toronto, Canada) was used as the DICOM workstation software. Standard RIS and HIS software also resided on the workstations.

Two different computer types were evaluated: a Pentium (Intel, Santa Clara, CA) 350-MHz with 64 MB RAM, and a Pentium 667 MHz with 128 MB RAM. A standardized typical 244-word CT dictation was dictated using the above hardware variables and the time for loading and saving the dictation was recorded. Fifty-three words of the dictation were derived from two macro statements. The remainder of the dictation was an original continuous speech dictation. A single radiologist (R.E.) performed all of the dictations after several practice trials. Dictations were performed with and without the DICOM workstation software running. Word recognition rate was also evaluated. Equipment cost information was collected.

RESULTS

Dictation was possible with the Power Scribe window minimized and the DICOM workstation software maximized. The system was found to have a word recognition accuracy of close to 100% for dictation with and without the DICOM viewer. Mean report dictation time decreased from 126.6 seconds to 122.2 seconds with faster computer hardware. No significant change in spelling errors was noted with different hardware. An average of one word per 244-word CT dictation was missed. The most commonly missed word was "for" in place of "or" from the following sentence: "If clinically indicated, a CT scan without contrast for MR scan may be helpful." The performance of the DICOM viewer was noticeably slower on the older computers. Hardware costs ranged from $3,200 to $5,300 depending on processor speed and additional monitors.

CONCLUSION

Better utilization of limited reading room space and computer hardware budgets may necessitate combining software applications. Speech recognition was excellent on all machines. Faster dictations were obtained with faster computer hardware. The eFilm DICOM viewer was noticeably slower on the older computer. Combining continuous speech recognition software with DICOM workstation software on the same platform is feasible and can lead to substantial savings of space and hardware cost. RIS, HIS, E-mail, and internet software were also helpful. Widespread implementation of a commercial continuous speech recognition system had occurred 2 years prior to this study. Soft-copy reading of radiology studies has been used on a limited basis for 3 years.

REFERENCES

1. Ratib O, Valentino DJ, McCoy MJ, et al: Computer-aided design and modeling of workstations and radiology reading rooms for the new millennium. Radiographics 20:1807-1816, 2000

2. Ramaswamy MR, Chaljub G, Esch O, et al: Continuous speech recognition in mr imaging reporting: advantages, disadvantages, and impact. AJR 174:617-622, 2000

3. Lemme PJ, Morin RL: The implementation of speech recognition in an electronic radiology practice. J Digit Imaging 13:153-154, 2000 (suppl)

4. Schwartz LH, Kijewski P, Hertogen H, et al: Voice recognition in radiology reporting. AJR 169:27-29, 1997

5. Rosenthal DI, Chew FS, Dupuy DE, et al: Computer-based speech recognition as a replacement for medical transcription. AJR 170:23-25, 1998

6. Collins J: Computer speech recognition for radiology dictation. SCAR News, vol 11, issue 3, July 1999