Experience With Voice Recognition in Surgical Pathology at a Large Academic Multi-Institutional Center

Hyunseok Peter Kang, MD,[1] S. Joseph Sirintrapun, MD,[2] Rick J. Nestler,[3] and Anil V. Parwani, MD, PhD[3]

Abstract

There are few reports of institutional use of voice recognition technology in clinical practice. We describe our experience with voice recognition–integrated synoptic-like dictation, associating templates with key spoken phrases, that we have used in gross examination of common specimens and as a major component of our workflow since 2001.

The primary application is VoiceOver Enterprise (Voicebrook, Lake Success, NY), which uses Dragon NaturallySpeaking Medical Edition (Nuance Communications, Burlington, MA) as its speech engine. This integrates with the anatomic pathology laboratory information system (APLIS) and other applications, such as Microsoft Office (Microsoft, Redmond, WA).

The largest user group, pathology assistants, mainly dictates biopsy reports, numbering approximately 210,000 specimens since 2001. The technology has been useful in our anatomic pathology workflow and provided a good return on investment, including marked improvements in turnaround time, results standardization, error reduction, and cost savings. The most helpful features of the software are templating, the seamless integration with APLIS, and the voice command creation tools.

Voice recognition technology has been available for more than a decade; however, widespread institutional use of this technology in clinical practice has been rather anecdotal. Numerous studies, mainly in radiology,[1-4] have compared the accuracy to human transcription. Despite the initial overhead setup cost, there is potential for considerable savings with transcription services and shorter turnaround time. The accuracy of voice recognition technology has been described as being as high as 99%,[4] whereas other studies have shown slightly lower accuracy than human transcription.[5] When extrapolated over large clinical reporting volumes, these small differences in accuracy lead to considerable editing and correction times. Rather than implementing this for all aspects of pathology reporting, some institutions have demonstrated applicability of voice recognition for dictation in scenarios amenable to standardized templates, such as autopsies and gross descriptions, using synoptic-like preprogrammed text associated with key descriptive spoken phrases.[6,7]

At the University of Pittsburgh, Pittsburgh, PA, we have been using anatomic pathology laboratory information system (APLIS)-integrated synoptic-like dictation in standard of practice with "grossing" of common specimens since 2001. We have found it useful from several aspects, such as reduction of errors and decrease in turnaround time. We report our experience with the use of this technology, which forms a major component of the workflow at our institute.

Materials and Methods

At the University of Pittsburgh, the sites of use include the 5 main hospitals, including Shadyside, Presbyterian, and Children's hospitals, the dental school, and 1 satellite hospital.

We have gone through several iterations of software: the current primary speech recognition and documentation application is VoiceOver Enterprise, version 4.01 (Voicebrook, Lake Success, NY), which uses the Dragon NaturallySpeaking 9.5 Medical Edition (Nuance Communications, Burlington, MA [formerly ScanSoft]) as its speech engine. New users undergo a 1-hour training session, and setting up a voice profile takes approximately 10 minutes. Dictation can be performed naturally at speeds of 160 words per minute. The vocabulary dictionary exceeds 260,000 words, with more than 60,000 specialized medical terms and phrases. The Voicebrook module integrates with the APLIS and other Windows desktop applications such as Microsoft Office (Microsoft, Redmond, WA). Hardware computing requirements are also reasonable, with most current personal computers in use being able to handle the application. The key feature of the module is the templating utility by which departmental reporting standards may be used, such as those used in gross examination reporting in biopsies and autopsies. Voice command creation tools, which can create shortcuts or automate routine steps in processing, are also very helpful. For example when a pathology assistant says "gross history" at a certain spot in the application, Voiceover will click the "Edit Text" button, which opens Microsoft Word, and the Clinical History template will automatically insert in the Clinical History text field.

To create reports using the voice recognition system, the user logs into VoiceOver and loads his or her speech file. He or she can then select any of the buttons in our APLIS by voice command. Once the patient's report is opened, saying "insert template" causes the Template Center window ❙Figure 1❙, which shows all of the user's templates, to display, from which the appropriate one can be selected and inserted. This can also be done by saying "insert <template name>." When the template is inserted into the report, the cursor is automatically placed into the first bracket [ ], and the user can dictate text into this bracket. Once the user is finished with that item, he or she says "next item," and the cursor is moved to the next bracket for dictation. There is also the option of directly dictating free text. When finished with that portion of the report, the user issues the command "return to CoPath" to minimize Microsoft Word and bring up the CoPath (Cerner, Kansas City, MO) window.

The Mann-Whitney U test was used to calculate statistical significance.

Results

Eighty user profiles, including 48 pathology assistants, 12 residents, and 20 attending physicians, have access to the Voicebrook module. At any point, 15 to 25 of them are actively using the program. Pathology assistants are the largest user group and dictate mainly biopsy reports. The program is also used for final diagnosis dictation in dermatopathology, dental pathology, and transplant pathology. Since 2001, the templates have been used for roughly 210,000 specimens. Before this time, the dictation was mostly free text with some paper-based templates that were read from.
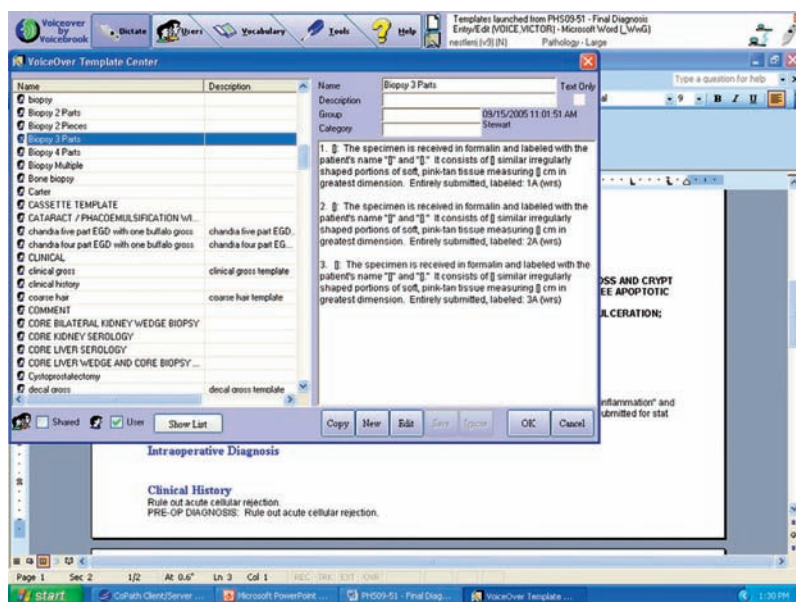

❙Figure 1❙ VoiceOver Template Center. When the user says "insert template," the VoiceOver template center window will display. This window shows all of the user's templates. The user can then select the template from the list and click or say "OK" to insert the template into the report.

**Comparison of Mean and Median Gross Examination
Completion Turnaround Time During 1 Month at 1 Hospital**

	Pre–Voice Recognition	Post–Voice Recognition
No. of specimens	1,109	937
Turnaround time (min)		
Mean	554.4	102.8
Median	203.5	30

There have been several benefits of using this technology. First, it has resulted in a marked reduction in turnaround time. The data for accession to gross completion turnaround time for one of the main hospitals in a representative month before and after implementation are shown in ❚**Table 1**❚ and ❚**Figure 2**❚, demonstrating that there was a statistically significant ($P < .001$), 81% decrease in average turnaround time, with the median turnaround time decreasing from more than 3 hours to 30 minutes. This trend is also conserved for final diagnosis dictation. ❚**Figure 3**❚ displays the data for 1 year on specimen turnaround time for one of our transplant pathologists. (Transplant pathology was used as an example because all but one of the pathologists in this area use voice recognition for final diagnosis dictation.) The percentage of cases signed out within 1 day improved by 89% ($P < .001$), even with a 68% increase in the case load. Second, the transcription error rate has fallen dramatically. ❚**Table 2**❚ provides the error rate in dictation and voice recognition over one year for the hospital represented in Figure 2, demonstrating that the total number of errors (potential adverse events caught before signing out the report) decreased by 48%. If we take into account the fact that a significant portion of the work is still processed through dictation, the effect is even more pronounced.

Discussion

Computerized speech recognition for radiology was first described in 1981.[8] Owing to limitations in hardware and difficulty of use (requiring pauses between words, training time of several hours), it remained essentially a curiosity for more than a decade. As computer systems have become more powerful while decreasing in cost and the software has matured, the use

**Comparison of Number of Transcription Errors (Potential
Adverse Events) During 1 Year at 1 Hospital**

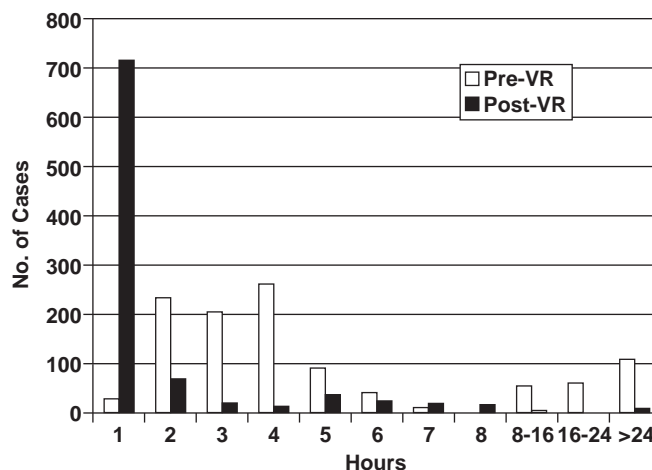| Adverse Event Setting	Pre–Voice Recognition	Post–Voice Recognition
Dictation	504	204
Voice recognition	—	56
Total	504	260

❚**Figure 2**❚ Comparison of gross examination completion turnaround times during 1 month at 1 hospital. There was a significant decrease in turnaround time, with gross descriptions for the majority of specimens being completed within 1 hour after the implementation of voice recognition (VR) technology.

of voice recognition systems has become much more widespread, especially in radiology reporting.[9] There have also been reports of implementation in other clinical specialties, such as orthopedics, pediatrics, and emergency medicine.[1,4,10-12]

The largest benefit that most of these studies have described is a decrease in turnaround time,[3,13-15] resulting in increased satisfaction from clinicians and administrators, as well as less time wasted fielding telephone calls for preliminary reports. Another advantage is that the reports are edited in real time (as opposed to a few days later when
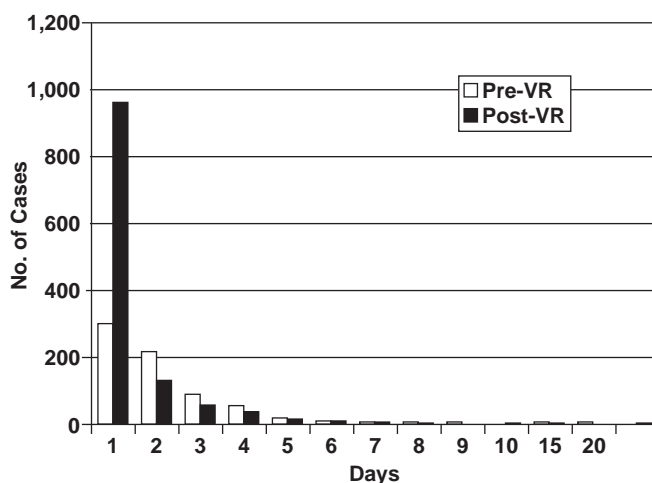

❚**Figure 3**❚ Comparison of specimen turnaround times during 1 year for a transplant pathologist. Post–voice recognition (VR), the number of cases signed out in less than 1 day increased more than 3-fold.

the transcriptionist completes them), leading to a reduction in major dictation and transcription errors.

The biggest disadvantage of using voice recognition technology is an increased editing burden on the part of physicians,[3,11] which is usually described as a major factor in the lack of widespread acceptance in pathology reporting.[5] Another limiting factor is that users need to modify their dictating behavior, formulating the report beforehand, to minimize "disfluencies" (hesitation, fragments, interruptions) that the software might interpret literally and cause a reduction in accuracy. This is partly because human transcriptionists perform higher-level functions, such as formatting and grammar correction, which computer-based systems cannot handle as efficiently as rote transcription.[16] The root cause of these reasons can be summarized as a lack of standardization in pathology reporting: there is variability in how reports are generated and their contents and more emphasis on free-text reporting rather than synoptic reporting.

Our experience at the University of Pittsburgh Medical Center is one of the few reported studies of successful use of voice recognition technology in clinical practice.[7,17] This technology, which converts speech to text in computer systems, is readily available in various software programs. The use of preprogrammed templates, to which gross descriptions of biopsy specimens are especially amenable, reduces the amount of text that needs to be entered and, thus, the impact of potential voice recognition errors, addressing one of the disadvantages of the technology by decreasing editing time.

Despite the initial startup cost, we have had good return on investment. By shortening turnaround time and standardizing reports, patient care has improved, and patient safety has been enhanced through reduction of transcription errors. The majority of this reduction reflects the decline in dictation errors in the gross and microscopic description text fields, which is most affected by the use of templates. We have found the most helpful feature of the software to be the templating utility and the seamless integration into the APLIS. Templated synoptic dictation provides voice recognition software the ability to "learn" from users through repetition. This leads to less editing and correcting of reports. Because of our success with the technology in gross workflow, we advocate widespread use of voice recognition technology in the setting of templated gross dictations. We have started expanding voice recognition into final diagnosis reporting and found that it also increases efficiency in this setting. Given that synoptic reports based on the College of American Pathologists cancer checklists are used for the majority of final diagnoses, we anticipate that voice recognition technology using synoptic templates will be ideal for this application.

From the Departments of ¹Pathology and Laboratory Medicine, Roswell Park Cancer Institute, Buffalo, NY; ²Pathology and Laboratory Medicine, Emory University, Atlanta, GA; and ³Pathology, University of Pittsburgh Medical Center, Pittsburgh, PA.

Address reprint requests to Dr Parwani: Division of Pathology Informatics, UPMC Shadyside Hospital, Dept of Pathology, 5230 Centre Ave, Room WG 02.10, Pittsburgh, PA 15232.

References

1. Ilgner J, Duwel P, Westhofen M. Free-text data entry by speech recognition software and its impact on clinical routine. *Ear Nose Throat J.* 2006;85:523-527.

2. Pezzullo JA, Tung GA, Rogg JM, et al. Voice recognition dictation: radiologist as transcriptionist. *J Digit Imaging.* 2008;21:384-389.

3. Ramaswamy MR, Chaljub G, Esch O, et al. Continuous speech recognition in MR imaging reporting: advantages, disadvantages, and impact. *AJR Am J Roentgenol.* 2000;174:617-622.

4. Zick RG, Olsen J. Voice recognition software versus a traditional transcription service for physician charting in the ED. *Am J Emerg Med.* 2001;19:295-298.

5. Al-Aynati MM, Chorneyko KA. Comparison of voice-automated transcription and human transcription in generating pathology reports. *Arch Pathol Lab Med.* 2003;127:721-725.

6. Klatt EC. Voice-activated dictation for autopsy pathology. *Comput Biol Med.* 1991;21:429-433.

7. Henricks WH, Roumina K, Skilton BE, et al. The utility and cost effectiveness of voice recognition technology in surgical pathology. *Mod Pathol.* 2002;15:565-571.

8. Leeming BW, Porter D, Jackson JD, et al. Computerized radiologic reporting with voice data-entry. *Radiology.* 1981;138:585-588.

9. White KS. Speech recognition implementation in radiology. *Pediatr Radiol.* 2005;35:841-846.

10. Parente R, Kock N, Sonsini J. An analysis of the implementation and impact of speech-recognition technology in the healthcare sector. *Perspect Health Inf Manag.* 2004;1:5. http://perspectives.ahima.org/index.php?option=com_content&view=article&id=78&Itemid=56. Published June 18, 2004. Accessed May 30, 2009.

11. Issenman RM, Jaffer IH. Use of voice recognition software in an outpatient pediatric specialty practice. *Pediatrics.* 2004;114:e290-e293. doi:10.1542/peds.2003-0724-L.

12. Borowitz SM. Computer-based speech recognition as an alternative to medical transcription. *J Am Med Inform Assoc.* 2001;8:101-102.

13. Chapman WW, Aronsky D, Fiszman M, et al. Contribution of a speech recognition system to a computerized pneumonia guideline in the emergency department. *Proc AMIA Symp.* 2000:131-135.

14. Kauppinen T, Koivikko MP, Ahovuo J. Improvement of report workflow and productivity using speech recognition: a follow-up study. *J Digit Imaging.* 2008;21:378-382.

15. Lemme PJ, Morin RL. The implementation of speech recognition in an electronic radiology practice. *J Digit Imaging.* 2000;13(2 suppl 1):153-154.

16. Schwartz LH, Kijewski P, Hertogen H, et al. Voice recognition in radiology reporting. *AJR Am J Roentgenol.* 1997;169:27-29.

17. Teplitz C, Cipriani M, Dicostanzo D, et al. Automated Speech-recognition Anatomic Pathology (ASAP) reporting. *Semin Diagn Pathol.* 1994;11:245-252.