Current State of Speech Recognition: Despite High Accuracy and Growing Interest, Behavioral Barriers Remain

by Donald T. Fallati

Speech recognition has moved beyond the experimental early-adopter stage, and while it has a way to go before it becomes truly widespread, the technology is actively deployed in hundreds of healthcare institutions. Adopters span the spectrum of inpatient and outpatient facilities, academic and community settings, and large and small organizations.

Speech recognition is also gaining depth and breadth within organizations. Radiology has been quick to embrace the benefits of speech recognition and now considers the technology a mainstream tool. Although no formal market penetration studies have been conducted to date, Dictaphone estimates that approximately 1,000 hospital radiology departments and outpatient imaging centers have already deployed the technology. The 2005 Healthcare Information and Management Systems Society Leadership Survey indicated that nearly 60 percent of respondents anticipate using speech recognition in the next two years.[1] Internationally, a recent Frost and Sullivan study predicted that "the annual growth rate is expected to reach double digits" for voice recognition over the next five years.[2]

There are two types of speech recognition technology. Both share similar barriers to adoption, as well as best practices for implementation.

Transcriptionist-assisted and Self-Editing

The first type of speech recognition technology is called transcriptionist-assisted, or background speech recognition. The other is known as self-editing.

Transcriptionist-assisted speech recognition fits comfortably into typical workflows, enabling transcriptionists to edit speech-recognized documents. It has gained momentum over the past few years due to advancements in hardware processing power and software developments that allow good recognition accuracy from dictation over the phone. A wide range of dictators who are phone-based can continue their normal patterns, with speech recognition working in the background to speed overall document production. For transcriptionists, software permits speech editing and typing to be done in the same environment, enabling them to switch work styles.

Self-editing is also gaining traction. Most radiologists using speech recognition complete documents themselves, and an increasing number of nonradiologists find the "once and done" approach beneficial because of the control they gain over their documentation. Managers seeking guidance on selecting motivated physicians to begin self-editing should look to some of the traditional high-dictation specialties such as orthopedics, cardiology, and neurology as well as to those that have shown a distinct inclination toward complete self-control of documentation. Caregivers in this category include mental health professionals, physical therapists, and hospitalists.

Measuring Success

Speech recognition promises transcription productivity improvements, so measuring the difference before and after implementation is important. On the transcriptionist-assisted side, benchmarks can be established for prespeech transcription productivity, and software then can measure the transcriptionist's throughput with speech recognition editing versus straight typing.

Productivity gains across HIM departments are frequently in the 40-50 percent range, while gains by individuals range from a few percent to well over 100 percent, according to estimates by Dictaphone. Analysis of data by prespeech productivity levels shows that some positive gains can be generated even for transcriptionists whose prespeech ratio of dictate-to-transcribe time

was as low as 2-2.5 to 1. Fifty percent output gains translate into a boost from 1,000 lines per day, for instance, to 1,500 lines. HIM departments are using such gains to attack persistent report backlogs, take on new work, and reduce or redeploy staff.

The physician self-edit style presents a different productivity measurement. Self-editing bypasses the transcription step, so the productivity gain is immediate and complete. Many organizations have realized significant savings in transcription costs through heavy reliance on self-editing.

Overcoming Barriers to Adoption

With high accuracy, demonstrated return on investment, and wide availability for many disciplines, speech recognition's barriers are not technological. The real challenge to its adoption remains behavioral change. To help overcome this, managers can devise ways to provide economic and noneconomic bonuses for those who achieve learning milestones and then demonstrate productivity gains for their departments.

On the physician side, the primary issue is the belief that reports will take longer to complete with speech recognition than with normal dictation. The key to adoption here is showing physicians that the technology both saves transcription time and reduces the amount of time spent dictating. This real savings in documentation time comes as a result of key features in self-edit speech recognition technologies, such as single trigger-word normal text blocks and voice-driven forms and templates.

Additional timesavers derive from the use of a rapidly emerging sister technology to speech recognition: natural language processing. Though still in development, this technology will extract clinical facts and structured sections from free-form narrative text. As a result, when physicians sit down at the computer to begin dictation, they can quickly query the software's database for clinical information such as a patient's allergies or problems. Time spent shuffling through a chart to check for the data will be eliminated.

Speech Recognition Comes into Its Own

Speech recognition continues to generate significant interest, and many healthcare organizations have either adopted or are seriously considering the technology. AHIMA has already convened one e-HIM® work group on speech recognition, and the topic figured prominently in the organization's 2005 "Scenarios and Solutions for the Future of Transcription" report, put forth in conjunction with the American Association of Medical Transcription.[1]

Why the interest? First, the drive to find cost and productivity improvements in all aspects of healthcare remains relentless. Most HIM departments in hospitals and outpatient groups struggle with rising report volumes, escalating costs, and the scarcity of sufficiently trained staff, all the while knowing that prompt report turnaround times contribute to better quality patient care. Allowing staff to work from home, paying hefty overtime costs, and outsourcing more work have helped, but not enough. Therefore, more and more managers are exploring technology options as a basis for sustained improvement.

The Promise of Speed and Automation in Narrative Documentation

A second reason for the growing interest in speech recognigtion is the strong momentum toward clinical automation and the broader electronic health record (EHR) movement. The clinical documentation process, which has changed little over the past 40 years, is being targeted for the same reasons that organizations are pushing adoption of computerized physician order entry systems, EHR software, and a host of other tools that seek to reduce error-prone and inefficient manual processes.

Speech recognition is part of the clinical automation discussion for a key reason. Dictation continues to be a popular and widespread tool for producing patient documentation. Its popularity rests on more than habit and convenience: dictation persists because narrative permits the detail, nuance, and comprehensiveness in patient reporting that caregivers demand. Highly structured documentation technology and "point-and-click" EHRs have met with physician resistance because

their benefits frequently do not outweigh their usability barriers. Speech recognition offers hope for many that speed and automation can at last be wedded to the comfort and benefits of narrative documentation.

Note

1. Fuller, Sandra R., and Jill Callahan Dennis. "Transcription's Future(s): AAMT and AHIMA Outline Scenarios for the Years Ahead." *Journal of AHIMA* 76, no. 7 (2005): 48-51.

Best Practices for Implementation

Speech recognition benefits from the same sound implementation advice as other clinical information systems. Securing the proper sponsorship of organizational leaders is critical to ensuring buy-in and shared objectives among key players, for example.

Beyond the common best practices, the most important factor for successful adoption of speech recognition is to plan for its use from the outset. Organizations should obtain sufficient dictator speech licenses to generate significant work, whether they use background or self-editing methods. This approach will help speed adoption and provide earlier, more meaningful evidence of the progress being made.

Transcriptionists will need to do a sufficient volume of speech recognition editing, typically at least half a day's work in the learning phase. When productivity gains are lagging, frequently transcriptionists are found to be editing only a few reports a day, an insufficient amount to create comfort and skill in using the tools. They should be discouraged from typing from scratch those jobs that can be speech-recognized and edited. In addition, speech-recognized jobs may need to be temporarily routed to transcriptionists who have been trained to edit; otherwise, this work will be typed and therefore useless as a training tool for speech recognition.

For physician self-editing, HIM professionals should create a time-saving environment. Managers can set up a starting library of report templates, normal text blocks, and forms. They can also consider leveraging natural language processing technology and setting user preferences to take advantage of the automatic text reuse timesavers. By having shortcuts ready, physician training and satisfaction can be significantly hastened.

Speech recognition is already delivering benefits to the HIM industry. As its use accelerates, those organizations and individuals who harness speech recognition's power to transform the clinical documentation process will realize the vision for a new e-HIM® future.

Notes

1. Healthcare Information and Management Systems Society. "2005 Leadership Survey." Available online at www.himss.org/2005survey.
2. Frost & Sullivan. "European Healthcare Voice Recognition Systems Markets." October 2005.

Donald T. Fallati *(don.fallati@dictaphone.com) is executive vice president of marketing for Dictaphone. He is also a board member on AHIMA's Foundation of Research and Education.*

Article citation:
Fallati, Donald T.. "Current State of Speech Recognition: Despite High Accuracy and Growing Interest, Behavioral Barriers Remain" *Journal of AHIMA* 77, no.4 (April 2006): 60,62.