

Speech Recognition Systems: Key Benefits to Healthcare



Technologies focusing on the generation, presentation and application of clinical information in healthcare, referred to as health informatics or eHealth solutions, have experienced substantial growth over the past 40 years. Pioneering studies have been published focusing on technologies for producing and using written or spoken text, known as computational linguistics, natural language processing, human language technologies, or text mining.

Health informatics or eHealth solutions enable clinical data to become potentially accessible through computer networks for the purposes of improving health outcomes for patients and creating efficiencies for health professionals. Language technologies hold the potential for making information easier to understand and access.

Speech recognition, in particular, presents some interesting applications. Speech recognition (SR) systems are composed of microphones that convert sound into electrical signals, sound cards that digitalise the electrical signals, and speech engine software that convert the data into text. Applications have been demonstrated in radiology resulting in a reduction in turnaround time of reports from 15.7 hours to 4.7 hours. Document processing within endocrinology and psychiatry also demonstrated improvements in productivity.

A study published in *BMC Medical Informatics and Decision Making* aimed to undertake a systematic review of the existing literature relating to SR technology and its applications within healthcare.

Material and Methodology

A systematic review of existing literature from the year 2000 was undertaken. The authors believed that only studies from 2000 onwards would use SR technology that was sufficiently accurate to be suitable for healthcare settings. Inclusion criteria were: all papers that referred to speech recognition (SR) in healthcare settings, used by health professionals (allied health, medicine, nursing, technical or support staff), with an evaluation of patient or staff outcomes. All research designs, both experimental and non-experimental, were included.

Six databases (CINAHL, EMBASE, MEDLINE including the Cochrane Database of Systematic Reviews, OVID Technologies, PreMED-LINE, PsycINFO) were searched by a qualified health librarian trained in systematic review searches, initially capturing 1,730 references. However, only 14 studies met the inclusion criteria and were retained. Of the fourteen studies retrieved, one was a randomised controlled trial; 10 were comparative experimental studies and most of the remaining were descriptive studies predominately using a survey design.

The studies were conducted in hospitals or other clinical settings including emergency, endocrinology, mental health, pathology, radiology, and dentistry departments. One study was carried out in a laboratory setting simulating an operating room.

The main outcome measures in the included studies were: productivity, including report turnaround time (RTT) or proportions of documents completed within a specified time period; and accuracy. The findings of the studies were heterogeneous in nature, with diverse outcome measures, which resulted in a narrative presentation of the studies.

Results

Productivity

Overall, most papers reported significant improvement in RTT with SR. Two studies reported a significant reduction of RTT when SR was used to generate patient notes in an emergency department (ED) setting and clinical notes in endocrinology. A longitudinal study (20,000 radiology examinations) indicated that using SR reduced RTTs by 81 percent, with reports available within one hour increasing from 26 percent to 58 percent. Similarly, the average RTT of surgical pathology reports was reduced from four days to three days with increases in the proportion of reports completed within one day (22 percent to 36 percent). Zick and Olsen reported that the reduction in RTT achieved by using SR in the ED resulted in annual savings of approximately \$334,000.

Results of another study reported significant differences in RTT between SR systems produced by different companies. The authors reported that Dragon software took the shortest time (12.2 minutes) to dictate a 938-word discharge report, followed by IBM and L&H.

Quality of Reports

The quality of the reports in seven studies was determined by comparing errors or accuracy rates. Taken together, the results from these studies suggest that human transcription is slightly more accurate than SR. The highest reported average accuracy rate across the included studies was 99.6 percent for human transcription compared to 98.5 percent for SR. However, an ED study found that reports generated by SR did not have

grammatical errors while typed reports contained spelling and punctuation mistakes.

Evidence from the included studies also suggests that error rates are dependent on the type of SR system. A comparison of three SR systems indicated that IBM ViaVoice 98 General Medical Vocabulary had the lowest overall error rates compared with Dragon Naturally Speaking Medical Suite and L&H Voice X-press for Medicine, General Medicine Edition, when used for generating medical record entries.

System Design

Some SR systems incorporated generic templates and dictation macros that included sections for specific assessment information such as chief complaints, history of present illness, past medical history, medications, allergies and physical examinations. Other researchers used SR systems with supplementary accessories for managing text information such as generic templates, medical or pathology terminology dictionaries, Radiology Information System (RIS) and Picture Archiving and Communication System (PACS).

Evidence from these studies suggests that the use of additional applications such as macros and templates can substantially improve turnaround times, accuracy and completeness of documents generated using SR.

Conclusions and Discussion

SR systems have substantial benefits for healthcare, but these benefits need to be considered in light of the cost of the SR system, training requirements, length of transcription task, potential use of macros and templates, and the presence of accented voices. The regularity of use enhances accuracy, although frustration can result in disengaging with the technology before large accuracy gains are made.

Expectations prior to implementation, combined with the need for prolonged engagement with the technology, are issues for management during the implementation phase. In most of the included studies, the reported error rates and improvements and other outcomes were achieved after only limited training was provided to participants who had no prior experience with SR. Training delivered varied from five minutes to six hours, but several researchers advised that either a pre-training period using any speech recognition system for one month or prolonged exposure with SR (one to three months) is preferred. This is confirmed by the improved turnaround times demonstrated in longitudinal studies.

The ubiquitous nature of SR systems within other social contexts will guarantee improvements in SR systems (software and hardware). The availability of applications such as macros, templates, and medical dictionaries will increase accuracy and improve user acceptance. These advances will ultimately increase the uptake of SR systems by diverse health and support staff working within a range of healthcare settings.

A thorough examination of the cost benefits of SR in specific clinical settings needs to be undertaken to confirm some of the economic outcomes proposed or demonstrated in this report.

Image Credit: Florida Technology Institute

Published on : Mon, 15 Dec 2014