



VCU

Virginia Commonwealth University
VCU Scholars Compass

Auctus: The Journal of Undergraduate Research and Creative Scholarship

2018

Speech Recognition Technology: Improving Speed and Accuracy of Emergency Medical Services Documentation to Protect Patients

Tan Tran

Virginia Commonwealth University

Follow this and additional works at: <https://scholarscompass.vcu.edu/auctus>

 Part of the [Health Information Technology Commons](#)

© The Author(s)

Downloaded from

<https://scholarscompass.vcu.edu/auctus/63>

This STEM is brought to you for free and open access by VCU Scholars Compass. It has been accepted for inclusion in Auctus: The Journal of Undergraduate Research and Creative Scholarship by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

Speech Recognition Technology: Improving Speed and Accuracy of Emergency Medical Services Documentation to Protect Patients

By Tan Tran

Abstract

Because hospital errors, such as mistakes in documentation, cause one in six deaths each year in the United States, the accuracy of health records in the emergency medical services (EMS) must be improved. One possible solution is to incorporate speech recognition (SR) software into current tools used by EMS first responders. The purpose of this research was to determine if SR software could increase the efficiency and accuracy of EMS documentation to improve the safety of patients of EMS. An initial review of the literature on the performance of current SR software demonstrated that this software was not 99% accurate, and therefore, errors in the medical documentation produced by the software could harm patients. The literature review also identified weaknesses of SR software that could be overcome so that the software would be accurate enough for use in EMS settings. These weaknesses included the inability to differentiate between similar phrases and the inability to filter out background noise. To find a solution, an analysis of natural language processing algorithms showed that the bag-of-words post processing algorithm has the ability to differentiate between similar phrases. This algorithm is best suited for SR applications because it is simple yet effective compared to machine learning algorithms that required a large amount of training data. The findings suggested that if these weaknesses of current SR software are solved, then the software would potentially increase the efficiency and accuracy of EMS documentation. Further studies should integrate the bag-of-words post processing method into SR software and field test its accuracy in EMS settings.

Speech Recognition Technology: Improving Speed and Accuracy of Emergency Medical Services Documentation to Protect Patients

Emergency medical services (EMS) is responsible for taking care of civilians in the United States when a health-related emergency problem occurs. The personnel who carry out operations to assist civilians in need of medical care are called first responders. The most recognizable aspect of the EMS is the ambulances used to transport personnel, such as first responders to and from emergency sites. However, EMS consists of more than just ambulances and first responders. The other units of EMS are the hospitals, the rehabilitation personnel, and the emergency department, to name a few. Overall, EMS is integrated into healthcare, public health, and public safety. The different parts of EMS and how they are run vary from state to federal levels (NHTSA's Office of EMS, 2017).

In the EMS, first responders in ambulances respond to a variety of health emergencies, such as strokes, heart attacks, skeletal injury, burns, etc. From every emergency, the first responders must keep a record of what occurred, and this is done using health records. Today, many EMS personnel use the electronic health records system, because the electronic records are easily integrated into the hospital system that the first responders bring patients to, and are not as unorganized as the handwritten method, which also prevents the need for paper (Rubin, 2016). However, with all documentation methods that require the recording of information, mistakes are prone to be made, especially in the hyper stressed environment that a first responder must operate in. The most common mistakes of documentation in the EMS prehospital setting are: undocumented patient initial condition during dispatch, inadequate narrative of patient before and after, insufficient detail of medical interventions, missing explanation for such interventions, and vague description of patient initial complaints (Page, Wolfberg & Wirth, 2016). In "Accuracy of EMS-Recorded Patient Demographic Data," Brice, Friend & Delbridge (2008) claimed that out of 360 prehospital records examined in the city of Pittsburgh, 301/360 or 83.6% had the correct patient name, 284/360 or 78.9% had the correct date of birth, and 120/360 or 33.33% had the correct social security number. Overall, the error rate for the just the demographics in all of the records was 73.9%. Because of these documentation errors, there are many possible harmful effects such as

loss of insurance coverage, loss of reimbursement to ambulance transport companies, and risk of further injury or death to the patient (Page, Wolfberg & Wirth, 2016). Hospital errors such as mistakes in documentation are predicted to cause one sixth of the deaths each year in the United States (James, 2013).

The high rate of errors in the prehospital setting are alarming and suggests that errors in documentation are a crucial problem that must be solved. One such solution to this problem is speech recognition (SR) software, as it has been widely used in many everyday applications, such as Google and iPhone. SR technology has greatly developed over many years and has been used in the healthcare setting for physician charting. In addition, computing processing power has increased, allowing for more effective processing of speech data, further improving speech recognition (Zhao, 2009).

In “Voice Recognition Software Versus a Traditional Transcription Service for Physician Charting in the ED,” Zick & Olsen (2001) claimed that SR was about as accurate as transcription of physician charting files, decreased the time it took to complete the health record and lowered the costs of creating such health records. Darcy, Gallagher, Moore, and Varghese (2016) claimed in their study, “The Potential of Using Voice Recognition in Patient Assessment Documentation in Mountain Rescue” that SR software has the potential to be utilized in emergency situations, as the use of SR in radiology was proved to be effective after the software was adapted to the health system.

Because SR software is not at 99% accuracy and still makes errors that may harm patients, SR software is not robust enough for immediate use in EMS settings.

However, because SR software can be improved by algorithms to increase accuracy by differentiating between similar phrases and filtering out background noise and decreasing interruptions in work flow for first responders in emergency situations, this software has potential for use in the future as a tool for filling out electronic medical records. Such an advance would improve the efficiency, accuracy, and safety for patients of the EMS services. Since there has not been substantial research on the adaptation of SR to prehospital emergency medicine, the current work aims to illustrate the ways in which SR can be adapted to work effectively in this setting.

A Standardized Measurement System for Accuracy Testing is Necessary

In order to limit the need for a medical transcriptionist who performs quality control on the sentences transcribed by the SR technology, SR must be as accurate as possible. In order to determine if this software is reliable enough, a standard system for determining accuracy or error types needs to be set for all researchers to compare to. In addition, the implementation of a SR system must be field-tested in order to ensure that this technology will improve, rather than cause more errors in the patients' electronic health record.

In "Doctors' Handwriting Gone Digital: An Ethical Assessment of Voice Recognition Technology in Medicine," Cheshire (2013) examined a 565-word document titled "A Christian Response to Adverse Outcomes Arising from Medical Error." In this study, the Nuance Dragon SR software achieved an 88% accuracy rate when subjected to the 565-word document. However, this software claimed "99% accuracy out of the box." Similarly, in "Evaluation of Google's Voice Recognition and Sentence Classification for Health Care Applications," Uddin, Huynh, Vidal, Taaffe, Fredenall, and Greenstein (2015) claimed that they studied the accuracy of a SR system in different configurations (different levels of training and different post processing techniques) being used in a perioperative services unit and found that the highest accuracy out of all the combinations of machine-learning techniques achieved less than an average of 84% accuracy. Uddin et al. (2015) stated that in the Google-only SR test, the 0-training had a median of 34% accuracy, the 5 training repetitions had 63% and the 10 training repetitions training had 69%. Uddin et al. claim that post-processing techniques also improved the accuracy of the SR system because the support vector machine technique yielded 82% accuracy with 5 repetitions of training and the maximum entropy post-processing technique yielded 84% accuracy with 5 repetitions of training. Both Cheshire (2013) and Uddin et al. (2015) found that the SR systems are nowhere near 99% accurate. In addition to Cheshire and Uddin et al., Groschel, Phillipp, Skonetzki, Genzwurker, Wetter, and Ellinger (2004) in "Automated Speech Recognition for Time Recording in Out-of-hospital Emergency Medicine—An Experimental Approach" studied the use of speech recognition in EMS settings by creating their own systems and found that the SR

systems they manufactured and tested gave accuracy percentages of 89% for a PC with headset, 85% for a pen-PC with headset, 84% for a PC with microphone and 80% for a pen-PC with microphone. When the systems were tested with emergency physicians in a simulated emergency scenario, the average accuracy percentage came out to be 75%, similar to Cheshire and Uddin et al.'s accuracy results. Because the SR software used by Cheshire claimed a 99% accuracy rate out of the box, a numerical standard can be set so that an acceptable SR system ready for the prehospital EMS setting must be at least 99% accurate for any field of text. This 99% accuracy rate standards for any software should be tested in a quiet room with minimal background noise to be a reliable constant for future researchers to test with. Currently, the testing of SR software from many studies have shown that this software is not close to 99%. Some may argue that 99% accuracy will be hard to reach, but if certain algorithms are used to help the software improve accuracy by fixing its weaknesses, then it will have a very good chance of reaching that 99%.

David, Chand, and Sankaranarayanan (2014) in "Error Rates in Physician Dictation: Quality Assurance and Medical Record Production," studied errors in 2,489 medical dictation files processed by 79 medical transcriptionists and SR software. David et al. (2014) stated that the study utilized standards from the Association for Healthcare Documentation Integrity, Medical Transcription Industry Association, and the American Health Information Management Association, which was different than Cheshire's study, where standards were not set using external guidelines from other entities. Similarly, in "Efficiency and Safety of Speech Recognition for Documentation in the Electronic Health Record," Hodgson, Magrabi, and Coiera (2017) studied the use of SR software to fill out electronic health records as opposed to using keyboard. Their study used an error ranking system which was similar to David et al.'s system but not Cheshire's, because Hodgson et al. based their error ranking system on a credible outside group - the United States Food and Drug Administration.

David et al.'s error ranking system consisted of two error types: critical errors and major errors, different from Cheshire's ranking system which consisted of four error types: misspellings that don't change the meaning, misspellings that change the

meaning but not understanding, misspellings that change the meaning of the text that could harm the patient, and misspellings that are offensive. David et al. defined critical errors as errors that compromise the safety of the patient such as include wrong patient citing, wrong drug, etc. and major errors were defined as errors that do not put the patient into risk, such as incorrect gender, incorrect age, etc. Hodgson et al.'s error ranking system had three categories: potential for patient harm, error type, and user error type. The potential for patient harm category consisted of labels of major, moderate, and minor impact on patient outcomes based on guidelines. Additionally, Hodgson et al. claimed that error type was split up into integration associated with technology, user error, and comprehension. Hodgson et al. claimed that user error types had 2 subcategories: omission or failing to complete a task and commission or errors that occurred when people executed a task incorrectly. This method of characterizing errors was more complex than Cheshire's and David et al.'s because there were multiple sublabels and errors could be categorized into more than one category. This analysis of the current research showed that methods of measuring errors from the SR software were not the same from each researcher, making it very hard to compare the data and the see how variables affect each SR system.

There must be a standard set of error documentation in order for SR systems to be utilized confidently in the field. That standard may include the 99% minimum accuracy requirement and an error ranking system, where the errors made by the SR technology will be ranked on a standard scale. Further research needs to be conducted on the various types of errors made in prehospital EMS settings so that a standard suited for this chaotic environment could be set for all future researchers to test with. When this standard is set, the researchers in the field of speech recognition use in healthcare can easily collaborate and compare results to one another so that SR software can easily be improved for use in the prehospital EMS setting.

Speech Disambiguation of Similar Phrases Can Improve Accuracy

SR software is unable to differentiate between phrases that are similar phonetically, and thus this is a source of error that changes the meaning of a sentence in a patient's electronic health record. Improving SR software through algorithms to detect subtle

variations in speech will negate this source of error and make SR software accessible to EMS first responders to document health records.

In “Doctors’ Handwriting Gone Digital: An Ethical Assessment of Voice Recognition Technology in Medicine,” Cheshire (2013) examined a 565-word document titled “A Christian Response to Adverse Outcomes Arising from Medical Error.” Cheshire stated that the program had an accuracy rate of 88%, with 65 total errors in the document. Cheshire claimed that the error types were 1 omission, 8 punctuation errors, and 46 phrase substitutions. Of the total errors, 14 errors changed the meaning of the phrases recorded but not the understanding of those phrases. Conversely, 47 errors changed the meaning of the phrases so that they were different from the intended phrase. Similar to Cheshire’s research, David, Chand, and Sankaranarayanan (2014) in “Error Rates in Physician Dictation: Quality Assurance and Medical Record Production,” studied errors in 2,489 medical dictation files processed by 79 medical transcriptionists and SR software. Similarly, in “Efficiency and Safety of Speech Recognition for Documentation in the Electronic Health Record,” Hodgson, Magrabi, and Coiera (2017) studied the use of SR software to fill out electronic health records as opposed to using keyboard and analyzed errors.

David et al.’s analysis of their errors was more in depth than Cheshire’s finding that medical transcriptionists can highlight 0.315 errors per dictation, while doctors make on average 315,000 errors in 1 million dictations. Additionally, David et al. claimed that the most common major errors were acronyms, gender mismatch, and age mismatch compared to phrase substitutions, the most common error found by Cheshire. David et al. reported that there was a statistically significant difference between the number of critical and major errors. They also found that the ratio of critical and major errors between transcription and speech recognition were significantly different, meaning that speech recognition and transcription created different types of errors. Cheshire stated some specific error examples such as: “couple ability” instead of the intended phrase “culpability,” “more Raleigh” instead of “morally,” “permission” instead of “omission,” “repayments” instead of “repentance,” “contraction” instead of “contrition,” “impression” instead of “confession,” and “approximately response” instead of “prompt

sympathetic response.” Similarly, David et al. included specific examples such as: “with arrest of the infant” as opposed to the intended, “followed easily by the rest of the infant” and “1000 units of Heparin” as opposed to “5000 units of Heparin.” Both Cheshire and David et al. agreed that these errors have a potential to cause medical error as well as misdirect the meaning of the message.

The common issue here is that SR is not sophisticated enough to differentiate between similar sounding phrases, and this could be compounded by the accents and speech variations of each voice. From the examples of incorrectly processed speech by the SR software, it can be concluded that the mistakes have the potential to physically harm the patient. Therefore, it is important to solve this issue for the SR software to be adapted for use in the prehospital EMS settings where medical terminology and numbers are mainly used. Such terminology includes many similar phrases, such as numbers used to quantify medical interventions, such as drug treatments or oxygen application. One way to differentiate between similar phrases is by using speech disambiguation through post processing algorithms.

The Bag-of-Words Technique

One such algorithm that will improve differentiation between similar words, is the bag-of-words post processing method. The bag-of-words post processing technique groups similar phrases defined manually by users together as matches. For example, if the user spoke “administer medications” into the SR software, but the software interpreted the input as “Minister medications,” the bag-or-words technique would check this and replace “minister medications” with “administer medications” if the user told the bag-of-words algorithm to match these similar phrases together. This method is best suited for large amounts data compared to other natural language processing algorithms that use machine learning, such as support vector machine. Machine learning algorithms often require a text corpus, or a group of words in order to train and base their processing methods off of.

In “Evaluation of Google’s Voice Recognition and Sentence Classification for Health Care Applications,” Uddin, Huynh, Vidal, Taaffe, Fredenall, and Greenstein (2015) used the bag-of-words post processing technique to modify the output by

Google's SR software. Uddin et al. reported that although the bag-of-words technique was simple, it significantly improved Google's SR software's accuracy. The simplicity of the bag-of-words technique is optimized for large data sets, such as the voice files that will be processed by the EMS first responders while filling out the electronic health records. Other methods, such as support vector machine, require training and more computing power than the bag-of-words method and because of this, not all of the computers used by the EMS first responders would be able to handle the machine learning methods with the large data that they have to process.

In "Challenges and Practical Approaches with Word Sense Disambiguation of Acronyms and Abbreviations in the Clinical Domain," Moon, McInnes, and Melton (2015) studied the use of machine learning and post processing techniques to solve word sense disambiguation of medical acronyms and abbreviations. Moon et al. claimed that simple techniques, such as bag-of-words, were effective for many cases in which machine learning was thought to be necessary, and this is similar to what Uddin et al. found in their research. Both Moon et al. and Uddin et al. state that the simplicity of bag-of-words worked well for these specific applications because the few number of phrases needed to be recognized. In "Towards Comprehensive Clinical Abbreviation Disambiguation Using Machine-Labeled Training Data" Finley, Pakhomov, McEwan, and Melton (2017) studied the use of machine learning to disambiguate abbreviations in medical texts and claimed that a large text corpus is necessary to train machine learning algorithms. Finley et al. stated that the use of a simplified corpus of abbreviations was viable for medical natural language processing systems. Similar to Moon et al. and Uddin et al., Finley et al. agreed that the bag-of-words technique was simple and reliable. Just from the bag-of-words, they were able to classify abbreviations with a 90% accuracy, an accuracy higher than other researchers in the field. The bag-of-words post-processing technique is better suited for use in the EMS prehospital settings because it is simple, requiring less computational power than the other machine learning techniques. If applied to SR software, the bag-of-words algorithm has the potential to solve the problem of misinterpreting? similar phrases by disambiguating the speech.

Filtering Background Noise

In the emergency response environment that EMT-First Responders operate in, many noises are present, including the wailing siren, the rattling of the ambulance, and the patient making noises. In order to achieve the highest possible accuracy, SR technologies must interpret these sounds and filter out the background noise.

Boll (1979) stated in “Suppression of Acoustic Noise in Speech Using Spectral Subtraction” that using a noise suppression algorithm called spectral subtraction can reduce the effects of acoustic noise added to speech. Mcaulay & Malpass (1980) similarly claimed in “Speech Enhancement Using a Soft-Decision Noise Suppression Filter” that breaking apart the spectral components of noisy speech could suppress the background noise altogether. Using a diagnostic rhyme test, Boll (1979) proved that the spectral subtraction method enhanced the quality of the voice when used to preprocess the voice data, which was similar to Mcaulay & Malpass’ (1980) results in which their algorithm that filtered the spectral components of the voice file produced “more pleasant-sounding speech to the ear since the annoying and tiresome background noise was removed.” In “Speech Enhancement Algorithm to Reduce the Effect of Background Noise in Mobile Phones,” Premananda & Uma (2013) stated that a psychoacoustic frequency method based on spectral algorithms was able to improve the quality and comprehension of the speech in noisy environments with amplified background noise. These results are very similar to those of both Boll and Mcaulay & Malpass, because they showed that a technique that used algorithms to analyze and optimize the spectral components of the speech signals were able to decrease background noise as well as improve the quality and clarity of the speech being filtered beforehand by the spectral algorithms. Additionally, Premananda & Uma (2013) claimed that the time domain technique was also able to minimize the effects of background noise on the speech signals. In “Combining the Evidences of Temporal and Spectral Enhancement Techniques for Improving the Performance of Indian Language Identification System in The Presence of Background Noise,” Polasi & Krishna (2015) claimed that recognition performances of speech were increased by 1-3% for various speech processing techniques when combined with the spectral preprocessing algorithm. When temporal

and spectral algorithms are combined, they increased the recognition of the system by 10-15%. The results from Polasi & Krishna's (2015) experiment supported the results of the experiments of Boll, Mcaulay & Malpass, and Premananda & Uma. From these studies, the conclusion can be drawn that many methods are currently being developed to digitally filter out background noise, primarily through analysis and rearrangement of the spectral components of the speech patterns. These results have shown that the processing of the spectral components of the speech patterns have improved the understanding of background noise, so when the modified speech inputs are run through the SR software, that software will have an easier time deciphering the words.

Other than algorithms and techniques to limit background noise, the positioning of the microphone can play a big role in increasing accuracy (Groschel et al., 2014). In "Continuous Speech Recognition for Clinicians," Zafar, Overhage, and McDonald (1999) stated that the optimal placement of the microphone to record speech from the user is one inch away from the mouth, as it helped them achieve an accuracy rate of up to 98% in their study. However, Groschel et al. stated that a microphone headset would not be ideal and suggested that a wireless collar microphone would be a good compromise. This positioning of the microphone would make the receiving voice clearer to the SR software, which would allow for better recognition in chaotic environments in conjunction with the spectral preprocessing techniques.

Fewer Interruptions

Using SR software and natural language processing to input data into the health record will prevent first responders from being interrupted as much as they are using an interface such as a computer with keyboard or a pen. This will decrease medical intervention error and improve patient outcomes.

In "Typed Versus Voice Recognition for Data Entry in Electronic Health Records: Emergency Physician Time Use and Interruptions," Dela Cruz, Shabosky, Albrecht, Clark, Milbrandt, Markwell, and Kedd (2014) discovered that the SR software statistically significantly lowered the number of interruptions compared to the keyboard input, with the keyboard causing interruptions at 5.33 times per hour while the SR causing interruptions at 3.47 times per hour. In "Efficiency and Safety of Speech Recognition for

Documentation in the Electronic Health Record,” Hodgson, Magrabi, and Coiera (2017) examined the impact of speech recognition technology vs traditional keyboard and mouse on the speed at which 35 emergency department physicians from three urban teaching hospitals in Sydney, Australia that were trained for speech recognition software filled out electronic health records as well as the number and severity of arising documentation errors using speech recognition software (Cerner Millennium suite with the FirstNet ED and Nuance Dragon Medical 360 Network Edition UK software) versus keyboard and mouse. They gave each physician eight of the same health documentation tasks using a commercial electronic health records system. The tasks included patient assignment, patient assessment, diagnosis, orders, and patient discharge and half of their tasks were complex and half were simple, the simple tasks only having two subtasks while the complex tasks having four subtasks.

Similar to Dela Cruz et al., Hodgson et al. studied interruptions by integrating interruptions into half of the health documentation tasks that physicians in their study had to complete. Hodgson et al. stated that the interruption was comprised of a single pop-up multiple choice question from an “Australian College for Emergency Medicine fellowship exam practice set.” Hodgson et al. reported that interruptions did not significantly affect the number of errors for the SR category, but did for keyboard and mouse, but cautioned that these results should not be used to generalize as interruptions are complex in nature. Dela Cruz et al. reported that SR did not change the time it took for physicians to chart compared to keyboard, but Hodgson et al.’s findings of interruptions did not match as Hodgson et al.’s research showed that speech recognition took 18.11% longer to finish tasks compared to keyboard.

SR by itself will cause fewer interruptions in work flow than keyboard input because with the keyboard, the user is frequently scanning the block of text for any mistakes that might have been included in the report. However, in terms of the time it took to complete the documentation task in a hospital setting, the research did not match and this was because interruptions have a multitude of components that could interact in complex ways. However, if an EMS first responder was working in the prehospital setting, interruptions would be very distracting, and therefore should be minimized

in order to prevent any extra sources of error, such as error from incorrect medical procedures.

Conclusion

In the EMS setting, errors are prone to occur in documentation, and that could cause loss of insurance coverage to the patient, loss of reimbursement to ambulance transport companies, and risk of further injury or death to the patient (Page, Wolfberg & Wirth, 2016). Hospital errors such as mistakes in documentation are predicted to cause one sixth of the deaths each year in the United States (James, 2013).

In order to alleviate this problem, SR software can be integrated into current EMS documentation in order to help create more accurate electronic health records at a faster pace. SR software has already been used in everyday life, from the software on your mobile device, to an automated telephone system used by some companies. Yet, current studies have shown that SR software is not at 99% accuracy as claimed by some manufacturers. Therefore, SR software is not robust enough for immediate use in EMS settings.

However, because SR software can be improved by algorithms to increase accuracy through speech disambiguation with the bag-of-words algorithm and filtering out background noise through spectral preprocessing techniques, and because SR software decreases interruptions in work flow for first responders in emergency situations, this software is a good candidate to use in the future as a tool to fill out electronic medical records in order to improve on the efficiency, accuracy, and safety for patients of the EMS services.

If current SR software could be optimized to work well under the conditions of the prehospital setting, then it would be a viable replacement for the current keyboard inputs that many EMS first responders use. The optimizations would include a spectral preprocessing algorithm to make the speech patterns more recognizable for the software, and the bag-of-words post processing method to distinguish between similar phrases. The greater implication of this study is if SR software is improved in the weak areas, then it has the potential to benefit patients of EMS services and would be good idea for EMS personnel to invest in SR software.

For future studies, more field tests of SR software are needed to get more data on the types of documentation errors that occur in the field of EMS due to the SR software so that a standard made specifically for SR software could be compiled for all researchers to use from now on. Additionally, researchers should incorporate the spectral pre-processing techniques and the bag-of-words post processing method into current SR software to get data on how this theoretical software would perform in real world scenarios such as in the prehospital setting and hospital setting, as well as in quiet areas where there is no background noise.

References

- Boll, S. (1979). Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on acoustics, speech, and signal processing*, 27(2), 113-120.
Retrieved from <http://ieeexplore.ieee.org>
- Brice, J. H., Friend, K. D., & Delbridge, T. R. (2008). Accuracy of EMS-recorded patient demographic data. *Prehospital Emergency Care*, 12(2), 187-191.
- Cheshire, W. P. (2013). Doctors' handwriting gone digital: An ethical assessment of voice recognition technology in medicine. *Ethics and Medicine*, 29(2), 71-76. Retrieved from <https://mayoclinic.pure.elsevier.com>
- Chittaro, L., Zuliani, F., & Carchietti, E. (2007). Mobile devices in emergency medical services: User evaluation of a PDA-based interface for ambulance run reporting. *Mobile Response*, 4458, 19-28. doi.org/10.1007/978-3-540-75668-2_3
- Darcy, R., Gallagher, P., Moore, S., & Varghese, D. (2016, June). The Potential of Using Voice Recognition in Patient Assessment Documentation in Mountain Rescue. In *Computer-Based Medical Systems (CBMS), 2016 IEEE 29th International Symposium on* (pp. 329-332). IEEE.
- David, G. C., Chand, D., & Sankaranarayanan, B. (2014). Error rates in physician dictation: Quality assurance and medical record production. *International Journal of Health Care Quality Assurance*, 27(2), 99-110. doi:10.1108/IJHCQA-06-2012-0056
- Dela Cruz, J. E., Shabosky, J. C., Albrecht, M., Clark, T. R., Milbrandt, J. C., Markwell, S. J., & Kegg, J.A. (2014). Typed versus voice recognition for data entry in electronic health records: Emergency physician time use and interruptions. *Western Journal of Emergency Medicine*, 15(4), 541-547. doi:10.5811/westjem.2014.3.19658
- Finley, G. P., Pakhomov, S., McEwan, R., & Melton, G. B. (2017). *Towards Comprehensive Clinical Abbreviation Disambiguation Using Machine-Labeled Training Data*. Paper presented at 2017 American Medical Informatics Association

- Symposium. Retrieved from <https://www.ncbi.nlm.nih.gov>
- Groschel, J., Philipp, F., Skonetzki, S., Genzwurker, H., Wetter, T., & Ellinger, K. (2004). Automated speech recognition for time recording in out-of-hospital emergency medicine—an experimental approach. *Resuscitation, 60*(2), 205-212. doi:10.1016/j.resuscitation.2003.10.006
- Hodgson, T., Magrabi, F., & Coiera, R. (2017). Efficiency and safety of speech recognition for documentation in the electronic health record. *Journal of the American Medical Informatics Association, 24*(6), 1127-1133. doi:10.1093/jamia/ocx073
- James, J. T. (2013). A new, evidence-based estimate of patient harms associated with hospital care. *Journal of Patient Safety, 9*(3), 122-128.
- McAulay, R., & Malpass, M. (1980). Speech enhancement using a soft-decision noise suppression filter. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 28*(2), 137-145. Retrieved from <http://ieeexplore.ieee.org>
- Moon, S., McInnes, B., & Melton, G. (2015). Challenges and practical approaches with word sense disambiguation of acronyms and abbreviations in the clinical domain. *Health Informatics Research, 21*(1), 35-42. doi.org/10.4258/hir.2015.21.1.35
- NHSTA's Office of EMS. (2017). *What is EMS?*. Retrieved from <https://www.ems.gov/whatisems.html>
- Page, Wolfberg & Wirth. (2016). *5 Costly EMS Documentation Mistakes*. Retrieved from <https://www.ems1.com/ems-products/ePCR-Electronic-Patient-Care-Reporting/articles/105246048-5-costly-EMS-documentation-mistakes/>
- Polasi, P. K., & Krishna, K. S. R. (2016). Combining the evidences of temporal and spectral enhancement techniques for improving the performance of Indian language identification system in the presence of background noise. *International Journal of Speech Technology, 19*(1), 75-85. doi.org/10.5121/ijwmn.2013.5114

- Premananda, B. S., & Uma, B. V. (2013). Speech enhancement algorithm to reduce the effect of background noise in mobile phones. *International Journal of Wireless & Mobile Networks*, 5(1), 177. doi.org/10.5121/ijwmn.2013.5114
- Rubin, M. (2016). *6 Questions EMS Leaders Must Answer About Buying an ePCR System*. Retrieved from: <https://www.ems1.com/ems-products/ePCR-Electronic-Patient-Care-Reporting/articles/112584048-6-questions-EMS-leaders-must-answer-about-buying-an-ePCR-system/>
- Uddin, M. M., Huynh, N., Vidal, J. M., Taaffe, K. K., Fredenall, L. D., & Greenstein, J. S. (2015). Evaluation of Google's voice recognition and sentence classification for health care applications. *Engineering Management Journal*, 27(3), 152-162. doi:10.1080/10429247.2015.1054752
- Zafar, A., Overhage, J. M., & McDonald, C. J. (1999). Continuous speech recognition for clinicians. *Journal of the American Medical Informatics Association*, 6(3), 195-204.
- Zhao, Y. (2009). Speech recognition technology in health care and special-needs assistance. *IEEE Signal Processing Magazine*, 26(3), 87-90.