Connecting the System to Enhance the Practitioner and Consumer Experience in Healthcare
E. Cummings et al. (Eds.)

doi:10.3233/978-1-61499-890-7-139

Using Voice-Activated Conversational Interfaces for Reporting Patient Safety Incidents: A Technical Feasibility and Pilot Usability Study

Owen SUN[a,1], Jessica CHEN[a] and Farah MAGRABI[a]
[a]*Centre for Health Informatics, Australian Institute of Health Innovation, Faculty of Medicine and Health Sciences, Macquarie University, Australia*

Abstract. Conversational interfaces and speech recognition capabilities are being increasingly used to create more natural and intuitive user interaction with digital technology. While voice-activated technologies have been used to support clinical documentation, their use for reporting patient safety incidents has not been previously investigated. The purpose of this paper is to assess the technical feasibility of an application for reporting incidents that combines a conversational interface with speech recognition software, and to undertake a pilot study of its usability. We built a prototype finite state-based application where incidents involving digital health technologies could be reported by answering five questions about the task being performed, the response of the software and the outcome of the clinical situation. Pilot tests showed that the conversational interface was usable. However, participants expressed concerns with speaking out loud about sensitive patient safety and quality improvement issues, including human error and system failures, in busy clinical environments. Further work is required to identify the clinical contexts in which conversational interfaces can be used to support incident reporting.

Keywords. Artificial Intelligence, Patient Safety, Conversational agent, Dialogue system

Introduction

Developments in technology have allowed for the use of software with conversational interfaces and speech recognition to increase workplace efficiency [1; 2]. Conversational interfaces are defined as "*computer systems with which humans interact on a turn-by-turn basis and in which spoken natural language plays an important part in the communication*" [3]. Recently, voice-activated interfaces such as Apple's Siri and Amazon's Alexa have integrated speech recognition with the conversational user interface to create a more sophisticated and interactive medium.

While voice-activated technologies have been trialled in healthcare environments to support clinical documentation with positive results, there are no previous studies investigating such applications for reporting patient safety incidents [4] [5]. The reporting of patient safety incidents is now widespread and is regarded as a "cornerstone"

[1] Corresponding Author.

of initiatives to improve the safety of health services [6]. Narrative reports by health professionals are a critical resource for understanding how and why incidents occur. However, not all incidents are reported and there is high variability in the incidents reported amongst different professional groups, with nurses more likely to report incidents than doctors [7]. A recent systematic review found the structure of incident reporting systems to be the second largest factor in both facilitating and preventing effective reporting of incidents [8]. Ineffective reporting systems are either time-consuming or overly complex and can discourage healthcare professionals from filing significant incidents and near miss events. Consequently, opportunities for learning and mitigation of patient safety risks are lost.

Voice-activated software offers some advantages for clinical documentation and has been linked to an improvement in healthcare workflow and proved to be an effective strategy in reducing operating costs [9] [10] [11]. Therefore, adopting a voice-activated conversational interface for incident reporting can reduce reporting times and simplify the reporting process.

Thus, the aims of this study are two-fold. Firstly, to assess the technical feasibility of an application for reporting incidents that combines a conversational interface with speech recognition software. Secondly, to undertake a pilot study of its usability for clinical contexts.

1. Method

1.1. Assessment of Technical Feasibility

A voice-activated conversational interface was built and published as a web application using JavaScript, HTML and CSS with a modified template provided by Tutorialzine [12]. The user interface was modelled on a previously tested online reporting system for patient safety incidents involving problems with digital health technologies [13]. Incidents could be reported by answering five questions about the task being performed, the response of the software and the outcome of the clinical situation (see Table 1). Users were presented with the choice of answering each question via keyboard entry, recorded speech entry, or both. Entries could be edited before saving the response and continuing (see Figure 1).

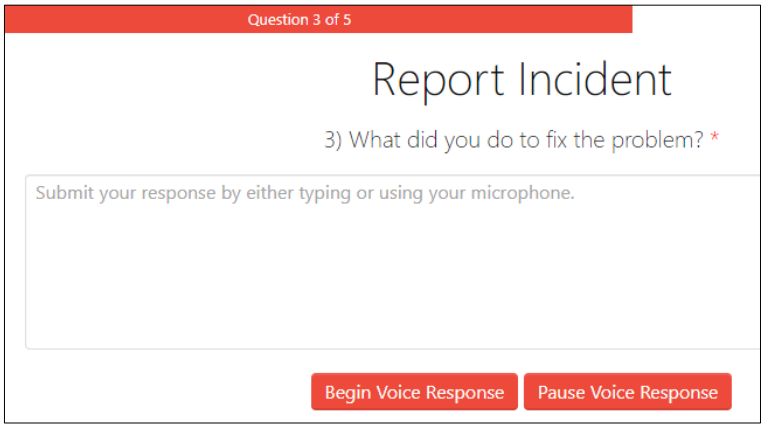

Figure 1. A screenshot of the web application with facilities for recorded speech entry.

At the backend, each response was stored as a text file in JavaScript Object Notation (JSON) on a database using MongoDB's database services (see Table 2) [14]. To transcribe speech, the 'SpeechRecognition' interface from Mozilla's 'Web Speech API' was used to access speech recognition services and event handling [15] . The application was published and deployed through 'Heroku', a cloud-based hosting platform [16] .

1.2. Pilot Study

The usability of the conversational-interface was tested with two scenarios that represented typical incidents reported by health professionals (Table 1). The design of these scenarios (one simple and one complex) was informed by a database of incidents from a previous study [13]. A specialist medical doctor and a pharmacist were asked to use the conversational interface to report the incidents (i.e. via the recorded speech entry) and asked to evaluate their experience. Testing was undertaken in a computer laboratory. Participant interactions with the system were observed and usability was measured with the System Usability Scale (SUS), a ten-item scale [17]. The SUS was completed via a paper questionnaire and participants were invited to provide any comments to support their responses in a face-to-face interview.

Table 1. Incident reporting scenarios used to test the usability of the conversational interface.

Scenario A	**Scenario B**
1. Describe the incident when you were using your computer and its peripherals (e.g. printer)? Tramadol is prescribed for a patient on a selective serotonin reuptake inhibitor (SSRI). *2. What was the result?* The patient developed symptoms of serotonin syndrome: sweating, tremor, confusion and was noted to be hypertensive, tachycardic and hyper-reflexive. *3. What did you do to fix the problem?* Advised the patient to stop intake of Tramadol immediately. Alternative SSRI compatible medication was considered. *4. Why did this incident happen?* The prescribing software did not provide any warning about the interaction. *5. How could the incident have been prevented?* The prescribing software should have provided a warning about the interaction.	*1. Describe the incident when you were using your computer and its peripherals (e.g. printer)?* A patient was visiting several doctors about severe abdominal pain. Nobody could diagnose the precise reason. Patient eventually was admitted to hospital where she was scanned and found to have a large mass on the kidney. She had a nephrectomy. When I went through her patient notes on the EMR I noticed that an abnormal scan had been filed in her patient file without appropriate follow up action by previous doctor. This abnormal scan was filed 5-6 months before her hospital surgery. *2. What was the result?* No action on abnormal scan sitting in patient file for 5-6 months. Patient later had emergency scan and surgery in a hospital (nephrectomy for mass on kidney). Patient is fine now. *3. What did you do to fix the problem?* I noticed abnormal scan sitting in her patient file. Paper records would have been much better and action could have been taken. A note on the abnormal scan was only displayed at the bottom of her EMR notes. Previous doctor obviously did not see it. *4. Why did this incident happen?* Abnormal scan results had not been noticed by doctors when looking in patient's electronic records. This may not have happened with paper records. *5. How could the incident have been prevented?* By having a formal system to track follow-up of test results.

2. Results

The voice-activated conversational interface for reporting patient safety incidents was successfully created and hosted online, allowing for responses to be made through a speech interface. Our application falls under the definition of a *Finite State-Based System*, in which interactions between the user and the application are a series of predetermined steps [18]. While the general purpose speech recognition service was largely adequate for the simple scenario, some medical terms in the complex scenario proved to be a challenge (e.g. nephrectomy). Common terms were also not recognised in some instances (e.g. patient was transcribed as "APT") (see Table 2).

Overall, participants found the system easy to use, quick to learn, and not unnecessarily complex (SUS scores: 97.5, 77.5). Participants' feedback regarding this was that the graphical user interface (GUI) was minimal, easy to use, and required minimal interaction. They also noted the good performance of a general purpose speech recognition capability for a health context without the need for specific training.

While the conversational interface was considered to be suitable for a quiet office or workstation environment, it was not seen to be practicable for other clinical contexts. The potential of background noise interfering with system use was viewed as a concern. Other problems were errors in reports due to unrecognised utterances, especially in settings where a health professional is in a shared work environment like a busy hospital ward or a community pharmacy. The lack of privacy when speaking to a reporting system was noted to be a major barrier for reporting sensitive details, including clinical errors and health system failures causing incidents. Other concerns were related to the speech system not recognising specific vocabulary and shorthand used by health professionals. For instance, pharmacists typically write and speak using abbreviations of Latin terms (e.g. "qid" is used instead of 4 times a day).

At the task level, participants commented that it was harder to use a conversational interface to report incidents because they needed to think and plan their response before speaking, as opposed to the ability to type while thinking about the sequence of events that led to the patient safety incident. The ability to go back and easily add details that were recalled in the process of typing up an incident report were seen to be an advantage of interaction via a keyboard and mouse.

Table 2. Scenario A as captured by the conversational interface.

Field	Example
Question 1	last night at about 11 o'clock APT on the mental ward was given Tramadol when it when there are some very high dose of ssris
Question 2	add round midnight the patient develops symptoms of serotonin syndrome we observe sweating tremor and confusion upon further investigation when I was that she was hypertensive tachycardic and hyperreflexive
Question 3	we're not as the patient to stop taking Tramadol in medially and removed it from her drug list without considering alturnative medications that are competitive to her ssris
Question 4	we believe this incident happens because the prescribing software do not provide warnings about as interaction
Question 5	this incident should have been prevented if the press writing software provided a warning

3. Discussion

We designed and tested a prototype application for reporting patient safety incidents via a conversational interface. In contrast to previous purpose built commercial applications [9], our study sought to apply a general purpose speech recognition service, demonstrating the ability to easily build conversational interfaces for healthcare applications using open source technologies.

While finite state-based dialog systems appear to be particularly suited for the well-defined task of reporting incidents, our study has uncovered several practical challenges. These challenges might prove to be a barrier for the use of conversational interfaces, particularly with respect to speaking out loud about sensitive patient safety and quality improvement issues including human error and health system failures in busy clinical environments. Many concerns also arose regarding the context of use. An example context is in a community pharmacy, where there are other pharmacists and patients talking in the same room. While transcribed speech entry can be potentially more efficient, noise interferences can cause unwanted words or conversations to be transcribed. This is a common concern with the use of any speech integrated system [19]. In purpose built clinical systems, this issue is addressed by training the system to prioritise known users above all others in the background. For example, an electronic medical records system in an Emergency Department is specifically trained to recognise and prioritise the voice of the clinician who has logged into the system.

Privacy is perhaps a bigger concern. Health professionals commonly work in shared environments and in front of patients, particularly in the hospital setting. It is in this environment where any person nearby can potentially hear the clinician during report entry and listen to another patient's information. Therefore, there is a need to consider the context in which the conversational interface will be used in terms of background noise and privacy.

Another important finding relates to the suitability of a conversational interface for the sporadic task of reporting incidents. We found that one participant had trouble formulating their response and entering it into system at the same time because they felt "too loaded". This is consistent with the literature outside healthcare where input and output using speech has been observed to disrupt problem solving and recall [20]. While specialist doctors may be well versed in regular tasks like the dictation of clinical notes, the irregular task of reporting incidents involves problem-solving and recall, requiring health professionals to recollect the sequence of events and identify problems that led to incidents. Further studies are needed to shed light on this observation as the impact of conversational interfaces on cognitive load has not yet been measured. It is thus not clear for which healthcare tasks conversational interfaces are suitable.

This study has some limitations. It used a prototype system to conduct a pilot usability study with a modest sample of two users. The interface was limited to the free-text elements of an incident report, we did not consider questions with structured response options. For example, incident reports typically ask reporters to indicate the nature of patient harm from a list of options (e.g. serious, major, moderate, minor and minimum). We also focussed on incidents involving digital health technologies but used generic open-ended questions so that the findings can be generalised to other patient safety domains. Another limitation was our use of a general purpose speech recognition service which was not specifically trained to handle medical terms. This may have increased the number of speech recognition errors.

4. Conclusion

While conversational interfaces are technically feasible and have the potential to be used for the reporting of patient safety incidents, our study has identified several socio-technical issues which show that they are not uniformly suited to the diverse contexts in which healthcare is delivered. As new technologies emerge there is a need for careful evaluation with respect to the nature of the task at hand and the many contexts of use. Further work is required to identify the clinical contexts within which conversational interfaces are particularly suited. The impact of conversational interfaces on cognitive load and any effects on the quality of reports submitted also needs to be examined.

Acknowledgements

We thank the pharmacist and medical doctor who assisted with pilot testing; Tony Wu and Vitaliy Kim for their assistance in programming the conversational-interface.

References

[1] E.C. Callaway, C.F. Sweet, E. Siegel et al., Speech recognition interface to a hospital information system using a self-designed visual basic program: Initial experience, *Journal of Digital Imaging* **15** (2002), 43-53.
[2] J. Chai, J. Lin, W. Zadrozny, Y. Ye et al., The Role of a Natural Language Conversational Interface in Online Sales: A Case Study, *International Journal of Speech Technology* **4** (2001).
[3] N. Fraser, *Assessment of interactive systems*, Mouton De Gruyter, 1997.
[4] M.P. Koivikko, T. Kauppinen, and J. Ahovuo, Improvement of Report Workflow and Productivity Using Speech Recognition—A Follow-up Study, *Journal of Digital Imaging* **21** (2008), pp 378-382.
[5] Y.D. Derman, T. Arenovich, J. Strauss, Speech recognition software and electronic psychiatric progress notes: physicians' ratings and preferences, *BMC Medical Informatics and Decision Making* 10 (2010), 44.
[6] J.C. Pham, S. Gianci, J. Battles, et. al., Establishing a global learning community for incident-reporting systems, *BMJ Quality & Safety* **19** (2010), 446-451.
[7] R. Lawton, Barriers to incident reporting in a healthcare system, *Quality and Safety in Health Care* **11** (2002), 15-18.
[8] S. Archer, L. Hull, T. Soukup, E. et al., Development of a theoretical framework of factors affecting patient safety incident reporting: a theoretical review of the literature, *BMJ Open* **7** (2017), 16.
[9] T. Hodgson, E. Coiera, Risks and benefits of speech recognition for clinical documentation: a systematic review, *Journal of the American Medical Informatics Association* **23** (2015), e169-e179.
[10] R. Parente, N. Kock, J. Sonsini, An Analysis of the Implementation and Impact of Speech-Recognition Technology in the Healthcare Sector, *Perspectives in Health Information Management* **1** (2004).
[11] S.L. Maree Johnson, Va Long, Paula Sanchez, Hanna Suominen, Jim Basilakis, Linda Dawson, A systematic review of speech recognition technology in health care., *BMC Medical Informatics and Decision Making* **14** (2014).
[12] D. Markov, Converting from Speech to Text with JavaScript, in: D. Markov, ed., 2017.
[13] F. Magrabi, S.T. Liaw, D. Arachi et al., Identifying patient safety problems associated with information technology in general practice: an analysis of incident reports, *BMJ Quality & Safety* **25** (2015), 870-880.
[14] mLab, MongoDB Hosting: Database-as-a-Service by mLab, https://mlab.com/, 2011.
[15] Mozilla , https://developer.mozilla.org/en-US/docs/Web/API/SpeechRecognition, 2018.
[16] Heroku, Heroku: Cloud Application Platform, https://www.heroku.com/, 2018.
[17] J. Brooke, SUS-A quick and dirty usability scale, Usability evaluation in industry, 189 (1996), 4-7.
[18] M.F. McTear, Spoken Dialogue Technology: Enabling the Conversational User Interface, *ACM Computing Surveys* **34** (2002), 90-169.
[19] T. Yoshioka, A. Sehr, M. Delcroix et al., Making Machines Understand Us in Reverberant Rooms, *IEEE Signal Processing Magazine* **29** (2012), 114.
[20] B. Shneiderman, The limits of speech recognition, *Communications of the ACM* **43** (2000), 63-65.