A WIRELESS SPEECH- AND TOUCH-BASED INTELLIGENT COMPREHENSIVE TRIAGE SUPPORT SYSTEM

P. Chang*, Y.H. Sheng**, D.W. Wang***, Y.S. Hsu**** and I.C. Hou*****

* Institution of Health Informatics and Decision Making, National Yang-Ming University, Taipei, Taiwan/ROC
** Industrial Research Technology Institute, Tainan, Taiwan/ROC
*** Institute of Information Science, Academia Sinica, Taipei, Taiwan, ROC
**** Emergency Department, Taipei Veterans General Hospital, Taipei, Taiwan/ROC
**** Department of Nursing, National Taiwan University Hospital, Taipei, Taiwan/ROC

polun@ym.edu.tw

Abstract: The primary objective of this study is to design a multimodal voice recognition (VR) support system for mobile nursing and to evaluate its feasibility. Systems were developed from scratch and evaluation was done in the Emergenct Department of a 2700-bed medical center. Results showed an average 99% of accuracy rate and subjects' average willingness to use of the system was 8.2, at a scale of 1-10 and 6 the baseline of acceptance. Our study demonstrated the values of VR integrated with other interfaces in mobile nursing.

Keywords: Voice Recognition, Mobile Nursing, Interface Design, Touchscreen, Comprehensive Triage

Introduction

Mobile computing in nursing could support and enhance nurses' point-of-care capability to assure patient's care quality and safety [1][2]. However, a better supporting information technology is still under investigation [1][3]. Tablet PC is heavy, fragile, and without satisfying battery life. The availability, maneuverability, and fire code factors make the wireless computers on wheels (COW) less attractive. Though many nurses have a PDA, the interface [3][4], fragility [4] and the availability of competing solution [5][6] limits its practically popularity.

Though the voice recognition (VR) technology has been expected in mobile healthcare settings [7][8][9], it has been the least studied and underutilized solution, compared to other interface solution, in healthcare [10]. Though many have studied its use in healthcare, most of them were for physician's dictation or transcription for physician's notes [7][9-21] and none was related to nursing.

The VR has advanced since the late 90's to achieve a better performance to have a better acceptance, be easier to implement with moderate IT software and hardware investment, and be more cost effective [7][9][10][13][14][17][18][21-24]. Though it has been estimated that the use of VR in healthcare will still keep growing [18][22], none prediction has put nurses' needs under consideration. The primary objective of this study is therefore to design a multimodal VR support system for mobile nursing and to evaluate its technical feasibility.

Materials and Methods

The study was done in the triage station of an emergency department (ED), which served around 300 new patients per day, of a 2700-bed medical center in 2003-4 in Taipei, Taiwan. There were a total of 72 nurses, including 20 certified triage nurses, in the ED and the majority of them had been trained with a well-developed wireless PDA triage support system in 2001-2 [5]. An IBM 3270 text-based triage information system, with fewer functions than those of the PDA system, has been operating in the triage station.

The same contents inside both the IBM and PDA systems, which composed of 203 items of the most popular chief complaints in the ED, 47 items of the demographic data, 20 items of the vital signs, 23 auxiliary items intended for easy data entry and the national triage guideline, were used to develop the new system. The contexts of IBM 3270 system were examined to assure the familiarity of system to users.

The system was designed with the voice and touchscreen modals. The voice modal was implemented under a wireless setting. The sample conversations with patients from experienced nurses in various shifts in triage were recorded to analyze their speech patterns and background noises. The screen layouts from the IBM 3270 and PDA systems were examined to design the screen layout of new system with a focus of facilitating the touchscreen operation.

The system was coded with VB 6.0. The Microsoft Speech SDK 5.1 and Simplified Chinese Language pack [25] were used to design the voice component. Since only the key-word, not free speech, type of information, such as chief complaints and vital signs, was documented in the current triage information system and no descriptive and free-text data were needed, just the command and control, not the dictation, function of Speech SDK 5.1 was used, and no extra dictionary was installed. The system was designed to read back the data recognized by the system from what the users just said through headset as a feedback by text to speech technique for confirmation [7]. Subjects could correct the mistake, if heard from the headset, by saying "Correction" and then repeat saying the correct term, or by clicking the touchscreen to enter the right term.

The system was designed with the object model in a 3-tier structure, as shown in Figure 1. The core object is the triage patient object which was designed in component technique to save all patient-specific data as the property value. The property value could be embedded with a simple rule such that an abnormal value could fire an alert reminder such as "high blood pressure attention" for 100 mmHg diastolic pressure value. The workflow agent is an intelligent module which could process simple preset procedures such as to guide process or inform physicians or nurses for emergency cases through sending email or short messages. The interaction coordinator module was used to control the speech, touchscreen, PDA [5] data upload and workflow modules. All data were saved in a XML file for voice recognition.

The intensive prototyping approach was used to develop the system. Prototype systems were presented to one senior experienced triage nurse for her feedbacks and comments for improvement. The system was put into formal evaluation after her thought it was ready.

The system was run under Windows-XP Sony PCV-LX900 PCs with P-III 1GHz CPU, 128 MB of RAM and 14" Pen-compatible LCD Tablet Display, which served as touchscreen for the study. Each PC was equipped with a wireless Logitech Cordless Freedom headset microphone with an effective range of 10 meters in our setting, which was wide enough to cover the triage area. There is a push button on the transmitter, with a theoretical capacity of 6 continuous working hours of life after half day charging, of the headset which is used to activate or deactivate the transmission.

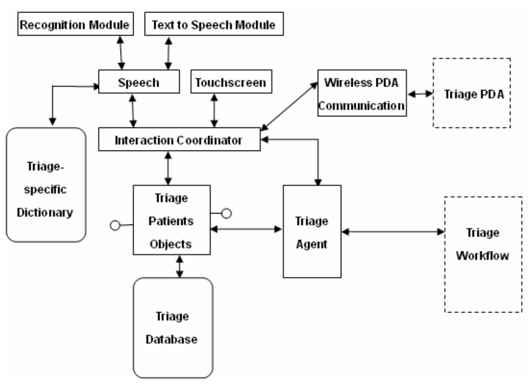

Figure 1: The systems structure which consisted of core components of triage patient object and coordinator object and modules of speech, touchscreen, PDA data upload, and workflow

Due to ED nurses' busy duty during our study, 30, out of 72, subjects were invited for the performance study based on their time availability. However, 10 more nurses were enrolled latter for the evaluation of willingness to use.

All subjects were first explained of the system. Then each one took around 30 minutes to run the voice training program and then, in a random order, took one out of three written scenarios, which were developed by experienced triage nurse for data entry and shown in Table 1. Scenarios were prearranged in a systematic way so each one could be equally tested by 10 subjects.

The data entry accuracy rates and time spent to finish a case were measured as the feasibility indicators. The accuracy rate was defined as the percentage of correct items entered over the standardized total items, which were predetermined by the senior triage nurse. The magnitude (dB) of background noise was recorded. Subjects were also asked to rate a single score after tests, at a scale of 1-10 and 6 is the baseline for acceptance, 10 the most preferred, to express their willingness to use of the system.

Table 1: Three written scenarios for subjects to test system with different interface mode

	Scenario
1	A 90 year-old female is brought in by 119 following a fall accident in bathroom last night. She complains severe pain of left upper leg which has obvious signs of fracture. She has history of hypertension and diabetes for more than ten years and she has no allergy history. The vital signs are BP=180/100 ; PR=78 ; RR=20 ; BT=36.7 in triage station. The patient is categorized to triage stage two and is transferred to trauma room with help.
2	Mr. Chang, a 72 year-old retired teacher comes to ED with the company of his family. He states that he is a healthy people with no chronic disease except allergy to penicillin and seafood. This morning, he feels dizziness; weakness; short of breathing and chest tightness suddenly. The vital signs are BP=70/40 ; PR=40 ; RR=20 ; BT=37 in triage station. The patient is categorized to the first priority and is transferred to resuscitation room with help immediately.
3	An 82 year-old male is transfer to ED by ambulance with the company of veterans home worker. The patient was found unconscious on the floor with cold sweating. He has DM history and with the symptoms and signs of drowsy and cold sweating. The vital signs are BP=120/70 ; PR=68 ; RR=17 ; BT=34.2 in triage station. The

patient is categorized to the first priority and is transferred to resuscitation room with help immediately.

ANOVA analysis was taken to examine subjects' performance difference among three test scenarios. Analysis was performed by use of the SAS 9.1 statistical program.

Results

The Touchscreen System

A vocabulary of 2-level 410 terms, under 15 categories such as vital sign, allergy, etc., was established to assure the performance of voice operation. 338 out of 410 terms were for the data entry and the others, for the control of system workflow among categories. The 2-level vocabulary was designed to support two steps of data entry: to say Category of contents first and then to say Content values to speed up the recognition response time and recognition accuracy.

The screen display was designed in larger forms, buttons and font size, under nurses' request. The screen layout was designed in 17 tap-able areas, as shown in Figure 2. There were three columns of tap-able areas. The left two columns show patient's data, such as "source," "way of arrival," "accompany," "chief complaints," "allergy," etc., and will turn to yellow when the corresponding data are entered, otherwise blue. While entering data, the auxiliary buttons for data entry would show up by side, as shown in Figures 3. The top-left corner, in pink, showed patient's basic data. The right column in orange is controlled by the system showing suggested triage acuity level, treatment area and category. The bottom at this column indicates the VR mode and volume level. The bottom row of the display is the status bar showing what data was just entered, with a button at right to end the system.

User could either manually tap the screen or say voice command to activate the system. When the system is operated by voice, the volume indicator at the right bottom will turn to green, otherwise, red. The majority of data entry can be done in no more than two taps under the touchscreen mode or two voice commands under the voice mode.

After the user ended the case by saying "Done," the system would suggest patient's acuity level and which examination room to deliver to for treatment.

The system was coded in an object model, composed of 6 collections, 61 property values, 47 methods and 18 events.

The Performance Evaluation

Subjects in this study were typically female, young, well educated and experienced. 85% were female. 85% were younger than 40. 75% were with BS or higher degree. 50% had worked in the ED more than 8 years. Almost all of them have ever worked in the triage. Before the test, 70% of them had ever heard of the VR but they didn't think they could be competent in using it. Interestingly, to the same question after the test, as high as 87% of them believed they could be very competent to work with the VR under little or even no training.

Though subjects tested with three different emergency scenarios, there is no evidence showing subjects' performance and the background noises among those scenarios were different, as shown in Table 2. Averagely, the accuracy rate for the system was as high as 99%; it took 108 seconds for subjects to finish one case; and the background noise was 61dB. Regarding subjects' willingness to use of the system, the average score was 8.2, with a standard deviation of 1.5.

Discussion

The study results demonstrate the practical value of VR in mobile nursing although we used only the command and control, not the natural language, technology of VR. We found no difficulty to set up an economic VR environment and to design a VR system to perform at a level of 95% average accuracy rate. Our results show that without specialized vocabulary the best accuracy rate was only around 75%. However, when equipped with a specialized, less than 500 terms, triage vocabulary, the accuracy became close to 95%, consistent with other study [7].

The issue of installing a special vocabulary or dictionary should not be a problem in nursing. The size of vocabulary for radiologists is close to 50,000[10], which is far larger than some vocabularies used in nursing such as ICNP® Version 1, in which the vocabulary has been downsized to no more than 800 terms [26]. Therefore, a compact set of vocabulary in nursing can make the use of VR very feasible [27].

Other factors such as accent and noise might not be significant in influencing the accuracy rate in our study. We did observe six nurses with particular accents and two nurses wearing masks working in real clinical settings but performed at similar level to others. One nurse even performed at a 100% accuracy rate. These nurses were all first time VR users and trained with same materials. The results also suggested that the ambient noise had no effect on the recognition accuracy, which is similar to the observations in the other study [13].

Interestingly, one tiny feature of microphone, which is the push button in the transmitter component, was capable of influencing the performance of system [13]. This is not because this tiny button could improve the system accuracy rate, but because it could reduce unnecessary errors by not responding to the background noise or other pointless or unimportant conversations. The on and off control by this button can control the system when to start or to stop processing and recognizing voice input.

Our study indicates that it is easy to work with VR systems. Though 70% of subjects regard themselves as novice VR system users before the test, about 90% of them believed they could be good at using the VR systems with minimum or no training after the test. During the test, the training time was only half hour. Therefore, an expectation of a zero-training voice-based system [7] might not be unreal.

Our subject ED nurses averagely took 2 minutes to finish coding of real but written scenarios under test environment. Though these written scenarios might be too simple to fulfill the work or need for a comprehensive triage [28], these did reflect the real needs for spot checks our triage nurses currently did, which took about 30 seconds to 2 minutes in general to finish a case.

The primary reason for the system to operate in a longer time is neither a long recognition response time nor a long error correction time. It is mainly because the system was designed to repeat back what nurses just entered by voice. This design is helpful to assure the data accuracy by reading back the data recognized by the system to nurses for confirmation but at a cost of doubling the system operating time. An useful alternative voice data confirmation mechanism should no doubt enhance the effectiveness of VR.

In our willingness to use survey, subjects showed very positive attitude and interest to use the multimodal system. Subjects' willingness to use was higher than the baseline willingness to use level. Subjects' free and easy switching to voice mode when they need their hands free for tasks like doing physical examination, measuring vital signs while simultaneously entering data inside triage station or in nearby, or switching back to touch mode when they prefer to enter data by quickly tapping the large buttons on the screen and to easily see and browse the results without bothering listening to the reading back of system no doubt can improve the system performance and attract users to use by providing a better and more flexible data entry interface.

Also notable, though the wireless setting is not part of technology for evaluation in this study, it should not be overlooked in its role to make the VR mode more useful. The mobility capability of entering data at a distance by VR will not be able to be illustrated if it is not used in the wireless setting. This condition has become common when more and more healthcare environments have equipped with wireless setting such as 802.11 wireless local area networks.

Better VR technology has kept progressing [7][9], therefore, if well integrated with other touchscreen devices in a wireless setting, the VR should be able to become an excellent solution for the mobile nursing in the future. It has been proposed that PDA combined with voice recognition technology could play the future role in mobile healthcare [4], and the performance of touch-based COWs could definitely be better improved with the VR [29][30].

Conclusions

Though this study was done in a test environment, we believe the multimodal voice-, and touch-based system should have the best chance to become the most useful and popular solution for mobile nursing in the near future. The multimodal VR system should be able to make the system perform at a level close to 100% accuracy rate and major concerns raised on its problem could be reduced to a negligible level in this multimodal mode.

Acknowledgments

This work could not be done without the active involvement and participation of the ED nurse members of the Taipei Veterans General Hospital.

References

[1] MALKARY G., (2005): Mobile computing in nursing study. Available from: URL: www.pdacortex.com/Mobile_Computing_Nursing_Study.htm (updated June 16, 2004), Accessed March 15, 2005.

[2] HUGHES SJ. ., (2000): Point-of-care information systems: state of the art. In: Ball MJ, Hannah KJ, Newbold SK, Douglas JV, eds. Nursing informatics: where caring and technology meet, 3rd ed. New York: Springer-Verlag, 2000; 242-251.

[3] MALKARY G, FARRELL A. ., (2005): Mobile nursing computing solutions: why the resistance to adopt? HIMSS 2005 Conference, Dallas Convention Center, Dallas, TX, USA, 2/13-17, Education Session #86.

[4] GILLESPIE G. ., (2005): PDA are willing, but will they be able? Health Data Management, December 2002. Accessed at onJuly 25, 2005. Available at http://www.healthdatamanagement.com/html/current/PastIssueStory.cfm?PostID=13550&PastMonth=December&PastYear=2002 .

[5] CHANG P, TZENG YM, WU SC, SANG YY, CHEN SS. ., (2003): Development and Comparison of User Acceptance of Advanced Comprehensive Triage PDA Support System with a Traditional Terminal Alternative System. In Mark Musen, Charles Friedman, Jonathan Teich (Editors), Biomedical and Health Informatics: From Foundations to Applications. American Medical Informatics Association 2003 Symposium, Marriott Wardman Park Hotel, Washington, DC, USA, 11/8-12, p140-144.

[6] RODRIGUEZ NJ, BORGES JA, SOLER Y, MURILLO V, COLON-RIVERA CR, SANDS DZ. ., (2003): PDA vs. Laptop: A Comparison of Two Versions of a Nursing Documentation Application, 16th IEEE Symposium on Computer-Based Medical Systems (CBMS 2003), New York, NY, USA, 6/26-27, p. 201-206.

[7] ERDEL T, CROOKS S., (2000): Speech recognition technology: an outlook for human-to-machine interaction. Journal of Healthcare Information Management, 2000;14(2):13-21.

ISSN: 1727-1983 © 2005 IFMBE

[8] MCCORMICK KA. ., (2000): Solving the vocabulary problem. In: Ball MJ, Hannah KJ, Newbold SK, Douglas JV, eds. Nursing informatics: where caring and technology meet, 3rd ed. New York: Springer-Verlag, 2000; 110-120.

[9] BRIGGS B. ., (2005): Speech recognition is finding its voice. Health Data Management, Februry 2004. Accessed on July 25, 2005. Available at http://www.healthdatamanagement.com/html/current/PastIssueStory.cfm?PostID=16913&PastMonth=February&PastYear=2004 .

[10] DEVINE EG, GAEHDE SA, CURTIS AC. ., (2000): Comparative evaluation of three continuous speech recognition software packages in the generation of medical reports. Journal of American Medical Informatics Association. 2000;7(5):462-468.

[11] GOEDERT J. ., (2005): Speech recognition technologies give voice to clinical data. Available from: URL: http://www.healthdatamanagement.com/HDMSearchResultsDetails.cfm?DID=13547 (updated Dec. 2002), Accessed March 15, 2005.

[12] ANONYMOUS. ., (2002): Speech recognition systems. Are they up to the task? Health Devices, 2002;31(2):65-71.

[13] ZAFAR A, OVERHAGE JM, MCDONALD CJ., (1999): Continuous speech recognition for clinicians. Journal of the American Medical Informatics Association, 1999;6(3):195-204.

[14] BAJAJ S. ., (2004): Expanding the reach of speech recognition. Extending speech recognition technology benefits to transcription workflow can boost productivity. Health Management Technology, 2004;25(3):16, 20-2.

[15] MOHR DN, TURNER DW, POND GR, ET AL. ., (2003): Speech recognition as a transcription aid: a randomized comparison with standard transcription. JAMIA, 2003;10(1):85-93.

[16] LUCAS MS. ., (2005): Speaking of efficiency. Health Management Technology, July 2001, accessed July 25, 2005, from http://www.healthmgttech.com/.

[17] DURLACH P. ., (2005): Talking up an emerging technology. Health Management Technology, April 2000, accessed July 25, 2005, from http://www.healthmgttech.com/.

[18] DREVENSTEDT GL, MCDONALD JC, DREVENSTEDT LW. ., (2005): The role of voice-activated technology in today's dental practice. Journal of American Dental Association. 2005;136(2):157-161.

[19] ISSENMAN RM, JAFFER IH. ., (2004): Use of voice recognition software in an outpatient pediatric specialty practice. Pediatrics. 2004;114(3):e290-293.

[20] AL-AYNATI MM, CHORNEYKO KA. ., (2003): Comparison of voice-automated transcription and human transcription in generating pathology reports. Archives of Pathology and Laboratory Medicine.2003;127(6):721-725.

[21] HENRICKS WH, ROUMINA K, SKILTON BE, OZAN DJ, GOSS GR., (2002): The utility and cost effectiveness of voice recognition technology in surgical pathology. Modern Pathology. 2002;15(5):565-571.

[22] VAN TERHEYDEN N. , (2005): Is speech recognition the Holy Grail? Available from: URL: http://www.healthmgttech.com/archives/0205/0205is_speech.htm (updated 2005), Accessed March 15, 2005.

[23] VON GREY M. , (2005): Technology moves into the mainstream. Health Management Technology, April 2000, accessed July 25, 2005, from http://www.healthmgttech.com/.

[24] ROSENTHAL DF, BOS JM, SOKOLOWSKI RA, MAYO JB, QUIGLEY KA, POWELL RA, TEEL M. , (1997): A voice-enabled, structured medical reporting system. Journal of American Medical Informatics Association. 1997;4(6):436-441.

[25] SPEECH SDK 51, (2002): for Windows applications. Available from: URL: http://www.microsoft.com/speech/download/sdk51/ (Updated June 24, 2002), Accessed March 15, 2005.

[26] ICNP, , (2005): accessed July 29 2005, available on http://www.icn.ch/icnp.htm.

[27] SPRING C. , (2003): Productivity Gains of Speech-Recognition Technology. Health Management Technology, January 2003, accessed July 25, 2005, from http://www.healthmgttech.com/.

[28] GILBOY N, TRAVERS D, WUERZ R. , (1999): Re-evaluating triage in the new millennium: a comprehensive look at the need for standardization and quality. Journal of Emergency Nursing. 1999;25(6):468-473.

[29] BENDER GT. , (1999): Touchscreen performance as a function of the duration of auditory feedback and target size. Unpublished doctoral dissertation, Wichita State University, 1999.

[30] DERON M. , (2005): How important is visual feedback when using a touch screen? Accessed July 25, 2005. URL: http://psychology.wichita.edu/surl/usabilitynews/2W/touchscreen.htm.

ISSN: 1727-1983

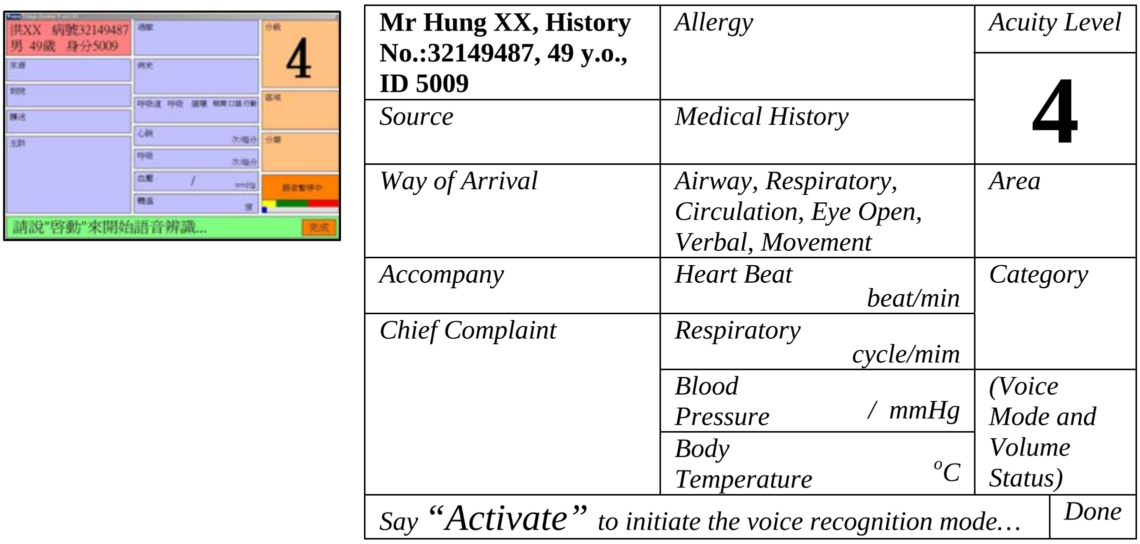

Figure 2: The snapshot of initial screen

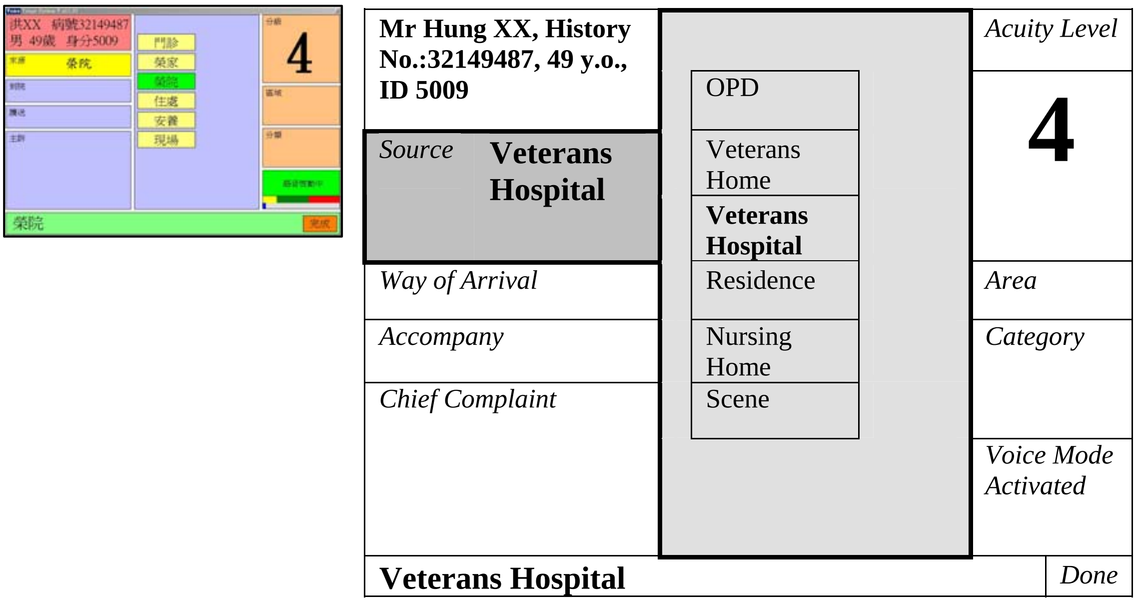

Figure 3: Under the voice mode, the voice mode button will turn to green. The screen shows that the user is telling the system the patient came from the Veterans hospital. The 6 buttons in light grey are the preset data items for users to tap under the touchscreen mode.

Table 2: The ANOVA analysis results of the performance and background noises for subjects (N=30)

	Test Scenario						p value	ALL	
	1 (N=10)		2 (N=10)		3 (N=10)			(N=30)	
	Mean	SD	Mean	SD	Mean	SD		Mean	SD
Accuracy	0.99	0.03	0.99	0.02	0.99	0.03	0.9565	0.99	0.03
Time to Finish	108.00	17.21	111.10	14.22	106.10	29.22	0.8688	108.40	20.59
Background Noise	59.25	2.37	61.17	7.32	61.13	3.96	0.6209	60.52	4.91

ISSN: 1727-1983