Voice User Interface Design for a Computer Aided Ureteroscopic Surgical System

Meng-Dar Shieh[1], Shih-Tyng Hsieh[2]
Department of Industrial Design, National Cheng Kung University, Tainan, Taiwan[1]
Yulon Group/ Hua-chuang Automobile Information Technical Center Co., Ltd.[2]
mdshieh2012@gmail.com[1], shihtyng@gmail.com[2]

Abstract. This paper presents a design for a computer aided surgical (CAS) system with a voice user interface (VUI) using Sami, a purpose-designed CAS system with VUI. Then, during Ureteroscopic surgery, we compare conventional UI -- as performed at the National Cheng Kung University Hospital (NCKUH) -- and the VUI of Sami. We invited three surgeons from NCKUH to participate in a Wizard of Oz test (WOZ test). Subsequently, the doctors were asked their opinion of Sami. At the end of the design process, the system framework of Sami was defined and programmed using Java. In the usability test, Sami demonstrated that the three surgeons could control the interface without using their hands, allowing them to see the nidus of the patient and administer the treatment. In the comparison results, the surgeons gave higher scores for convenience and speed to the VUI than to the conventional UI.

Keywords: Urology, Ureteroscopic surgery, Voice User Interface, Computer Aided Surgery, Design, Usability

1 Introduction

The keen competition of markets and the increasing demand for various products has pushed the development of industrial design far beyond the functions and costs of production. We must now consider the physical factors of a product, including its esthetics. Maslow's Hierarchy of Needs suggests that they can be categorized into several classes. Many of these needs are instinctive. Since User-Centered and User-Oriented concepts are vigorously promoted these days (Donald A. Norman, 1986), consumer focus groups have been playing a greater role in an effort to ensure that an upcoming product or service is successful. Difficulty in gaining consumers' confidence increases as market competition becomes more intense. Consequently, it is essential to encapsulate consumers' preferences in order to predict their purchasing choices. Researchers have used many methods in an attempt to predict how a product may appeal to a consumer in the marketplace.

In Taiwan, the medical devices used in operating rooms (ORs) are still operated using conventional physical methods. Since surgery requires a bacteria-free environment, and surgeons only have two hands, operating medical devices needs several surgical assistants. However, some problems do occur from time to time. For example, poor assignments allocation, divergent opinions among the assistants, and the assistants' personal mistakes. These factors reduce the efficiency and success rate of surgery. Consequently, this research attempts to understand and consolidate all the factors through observation and interview. By this method, we propose a VUI of a CAS system that is custom designed to cater to the demands of medical professionals here in Taiwan. After the usability test, there is an assessment to determine whether the designed VUI is suitable for the medical environment in Taiwan and whether it provides a progressive solution. This will enable doctors in the future realize so-called solo surgery. Eventually, it is predicted that the results of this research will provide a helpful option when constructing such a VUI of CAS system here in Taiwan.

It is likely that, in the near future, a single surgeon will be able to control several computer-based processes during a surgical procedure (Grange, Fong, & Baur, 2004). Studies in Taiwan and overseas still lack sufficient reference regarding VUI's suitability for ureteroscopy. Hence, the main objective of this study is to design a VUI in a CAS system (Sami) that is suitable for medical application. During the research process, the entire procedure of Ureteroscopy at NCKUH was closely examined in order to design a CAS system fit for ureteroscopy. The scenario is consolidated as an integrated description and separated into different parts, from which requirements are chosen and formed into a corresponding table. The surgeons who participated in the test rated every requirement on its importance, which subsequently became the basis of the VUI design.

After producing the prototype of the CAS system, which included VUI control, the three surgeons were invited to take the usability test, after which the results were analyzed. The VUI system's performance was based on five criteria: error, safety, cognitive load, satisfaction, and comparison. From the results, an assessment was made on whether the medical environment in Taiwan suits using VUI. This study provides advice on designing a VUI for a CAS of the future.

Voice User Interface

The Siemens Integrated OR System (**SIOS**, Siemens Medical Solutions, Siemens AG) integrates essential functions in the OR using a universal interface (CAN-open BUS system). The integrated operation system of the SIOS allows the use of different devices situated in both sterile and unsterile areas of the OR. Medical devices or OR components can be controlled also. All components are integrated as a unit and can be installed into a carriage device (Perrakis, Hohenberger, & Horbach, 2013).

OR1 is a fully functioning, integrated multi-speciality surgical suite intended for a hospital's medical information system (MIS). The central components of the OR1 are the Storz Communication Bus (SCB) and the advanced image and data archiving system (Aida) from Karl Storz, Tuttlingen, Germany. Both components allow monitoring, access and networking of the MIS equipment and other OR facilities, as well as the acquisition, storage and display of image, patient, and equipment data during the endoscopic procedure (Irion & Novak, 2000).

HERMES Control Center is an OR networking technology developed by Computer Motion, Inc. This system provides simplified control over devices in the OR, either by the surgeon or by the sterile field OR nurse. For example, the surgeon can increase the brightness of the endoscopic light source by speaking the commands 'HERMES ... light source ... brighter'. The scrub nurse can start recording with the video cassette recorder (VCR) by pushing the 'VCR' button on a touch-sensitive screen, then touching the 'Record' button that appears on the display. The HERMES system provides this control via a central controller for medical devices in the OR. Many HERMES-compatible devices may be attached to the network (Roe & Wang, 2000).

Perrakis et al. (2013) evaluated the VUI usability of two CAS systems: SIOS and OR1. They invited seventy-four surgeons from five nations to conduct the experiments and came to the following conclusions: The SIOS voice control was more effective and more reliable than that of the OR1. Several functional errors occurred with the OR1.

Wizard of Oz Test

In the field of human-computer interaction, a Wizard of Oz test (WOZ test) is a research experiment in which subjects interact with a computer system that they believe to be autonomous, but which is actually operated or partially operated by an unseen human being. The WOZ test is first mentioned by Dr. John Kelly in his dissertation at John Hopkins University circa 1980 and gives an experiment's participants the impression that they are interacting with a program that understands English as well as an educated native English speaker. In fact, at least in the earlier stages of development, the program barely limped along, only partly fulfilling its intended function. The experimenter, acting as "Wizard", surreptitiously intercepts communications between the participant and the program, and then supplies answers and new inputs as needed.

2. Research Framework

2.1 Requirement Analysis (Scenario)

Before surgery starts, the circulating nurses prepare instruments for the OR. During this time, the surgeon clicks on, reads, and checks through the patient's information in the patient file system. If the patient has any questions, the questions are answered by the surgeon. The patient is then placed on the OR table. A nurse anesthesiologist attaches the necessary instruments for the vital sign measurement. Then, the anesthesiologist executes the procedure of spinal anesthesia. After checking to see whether the patient is properly anesthetized, the surgery begins. The patient is positioned in the lithotomy position. The surgeon starts to set the OR table and wraps the patient's legs with an ace bandage. Then, the surgeon carries out the antiseptic procedures and places drapes over the patient body. A scrub nurse directs the circulating nurse to operate the image system by opening the patient's information file using a mouse. At the same time, the other circulating nurse adds a bottle of normal saline to the basin to help the scrub nurse prepare the surgical instrument. A circulating nurse announces the start time of the surgery and records the process.

A circulating nurse turns off the light of the OR before the endoscope is inserted into the urethra. The surgeon holds the endoscope and controls the direction. Meanwhile, the surgeon verbally records his progress: "left, going left… Oh! Here it is!" The surgeon considers the video on the screen to be too dark to see so he asks the circulating nurse to operate the button on the light source until the video on the screen is clear enough. After the circulating nurse double checks whether the light level is correct, she goes to the other end of the OR to get the guide wire for the scrub nurse. The surgeon takes the guide wire from the scrub nurse and inserts it into the endoscope. At the same time, a circulating nurse turns on the surgical lamp so that the instruments on the surgical cart can be seen clearly. Fortunately, the surgeon finds the calculi (kidney stones) and uses a foot pedal to take several pictures, which are saved digitally by an image system. The surgeon views the video input and indicates that the calculi should be removed immediately.

First, the laser lithotripter is moved by a circulating nurse to the OR room. The surgeon indicates where the laser lithotripter should be placed with his forefinger. Then, the laser conducting wire is passed from the circulating nurse to the scrub nurse. Before starting laser therapy, the laser conducting wire is placed carefully into the endoscope by the surgeon with the help of the scrub nurse. The circulating nurse switches on the laser lithotripter and the laser therapy begins. The surgeon moves the endoscope just in front of the calculi and aims at it. Using the foot pedal of the laser lithotripter, the surgeon begins to break up the calculi piece by piece. The surgeon encounters a problem: the calculi move away from the endoscope in the ureter when he fires the laser. He directs the circulating nurse to go to the other side of the OR table and push the button that tilts it to a head elevated position. There are brief interruptions in the process when the surgeon instructs the circulating nurse to adjust the OR table several time. After the laser therapy is finished, the nurses turn off the laser lithotripter and clean the instruments.

The surgeon looks at the video on the screen and decides that a double-J stent should be used for the patient. Finally, as the surgery seems to be completely successful the OR table is returned to its level position. To check whether the double-J stent is well placed or not, all medical staff (except those with protective wear) leave the OR and several X-rays are taken and saved.

The patient is returned to the ward. The nurses disconnect the endoscopes and gather up all the instruments. Meanwhile, the surgeon views the pictures and selects the important ones. He also types down in the surgical record system a description of what he has done during the surgery. The pictures and description are saved in two ways: one is a digital version for the patient file system; the other is a paper version, which is printed out, stamped, and saved with the other physical documents in the patient's file. The whole surgery takes 30 minutes and six medical staff: two surgeons, a scrub nurse, a circulating nurse, an anesthesiologist, and an anesthesiology nurse.

Table 1. System requirements table

Scenario	Requirements from surgeon	Requirements from observation
Prepare for surgery before patient arrives		Open and check the patient's information in the file system
Prepare for surgery (Patient arrives at OR)		Switch on the endoscope system
Surgery begin	Time out	Open the patient file in the image system
		Switch off the OR light
		Record the surgery information: time execution, blood loss, antibiotic, etc.
During the surgery	Record the note	Adjust the OR table
	Operate the endoscope	Answer the phone calls
	Search for the CT pictures	Switch on the surgical lamp
	Read out the literal report	
	Note the relevant pathological section	

	Call for medical machines	
Take pictures with endoscope	Take pictures	Adjust the light source
	Save pictures	
Medical therapy: laser lithotripsy	Change laser lithotripter status: standby, ready, off	
	Emit laser	
	Adjust the parameter of the laser	
X-ray	Operate the C-arm	
	Make X-ray films	
	Save X-ray films	
End of the Surgery		Announce the time elapsed
		Announce information about the next surgery: name of patient and medical staff
Edit the patient file	Type a report on the surgery: what was found, what transpired.	

2.2 Design Objects (Screening Result)

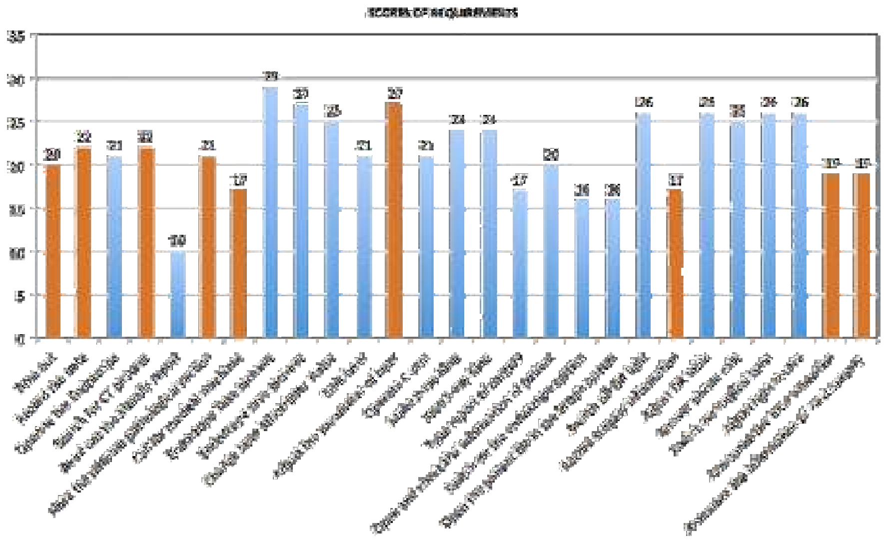

Fig. 1. Rating scores of the different procedures

We rated the importance of the different procedures on their scores (>=26): endoscope pictures (29), saving endoscope pictures (27), adjusting the parameters of the laser (27), switching off OR light (26), adjusting OR table (26), switching on surgical lamp (26), and adjusting the light source (26). Then, these important requirements were classified based on whether they had medical devices in common. Using this classification, six design tasks were defined: OR light, endoscope system, light source, surgical lamp, laser lithotripter, and, OR table.

VUI Design of SAMI (Wizard of Oz Test)

To test the user's voice commands and the usability issues of this system, we used Wizard of Oz (WOZ) as a host in this research. The WOZ test is often referred to as being a "discount" usability method equivalent to prototype testing, except no coding is required (Kortum, 2008).

Arrangements of WOZ Test

The purposes of the WOZ test is to resolve voice command and satisfaction issues: acceptability of functions, earcons, and, feedback voices. The WOZ prototype contains sound files and 3D rendering figures of medical devices used in the OR. Sound files of feedback voices are supported by Google while arcons (an Integral Field Spectrograph) are supported by Apple's iLife. The content of feedback voices has been defined by revising Sami. Cinema 4D rendered 3D figures of controlled objects. All medical devices and their positions in the OR were simulated.

Results:

Three surgeons from NCKUH were invited to participate in the WOZ test. All three had experience in ureteroscopy: 29 years, 5 years, and 1 year. The satisfaction issues are summarized as follows:

- Light source: the surgeons could not regulate the precise degree of brightness; therefore, the assigning brightness function could be dispensed with. However, the lightest and darkest functions proved to be more useful.
- Laser lithotripter: the surgeon preferred to get the parameter information before assigning the parameters.
- The feedback voices feature supported by Google enunciated clearly and the surgeons found it easy to understand.

3 Final Design

3.1 Switch on: Endoscope system/ Light source/ OR table

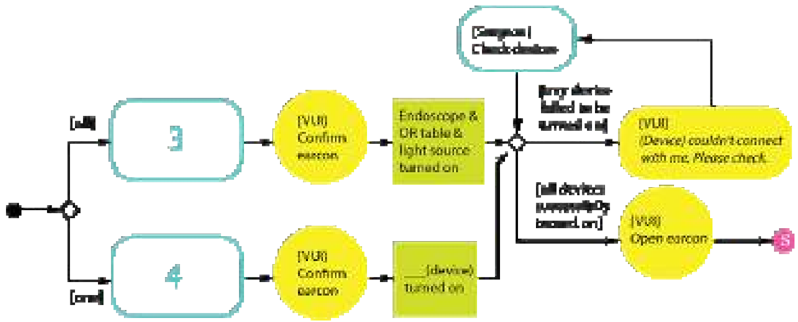

Fig. 4. The Final design of switch on function

Table 3. Voice commands and keywords of switch on function

#	Voice command	Keyword
3	**Operation start**.	▪ Operation && (Start‖On)
	Whole system powers **on**.	▪ Whole system && (Start‖On)
	Let's go.	▪ Let's go
	Start.	▪ Start
	Time out.	▪ Time out
4	Turn **on** the _____**(objects)**.	▪ (Endoscope ‖ Light source ‖OR table) && On
	_____**(objects)**.	
	Please turn **on** the _____**(objects)**.	▪ Endoscope ‖ Light source ‖OR table

3.2 Switch off: Endoscope system / Light source / Laser lithotripter/ OR table

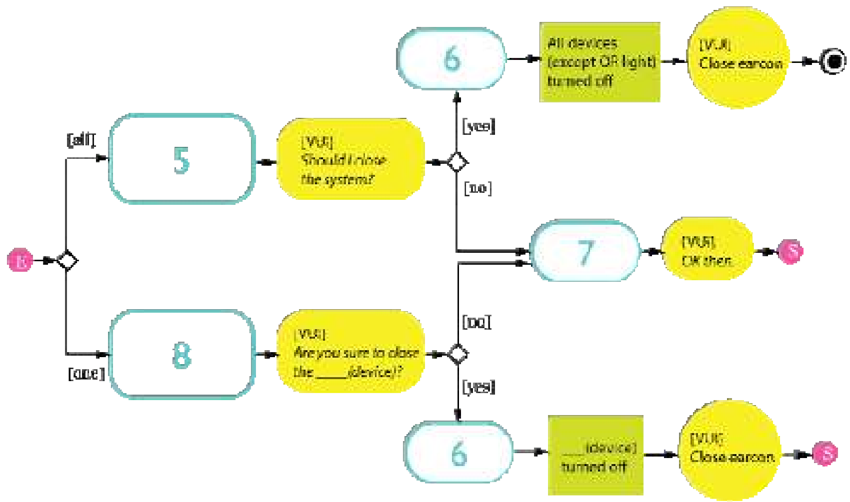

Fig. 5. The Final design of switch off function

Table 4. Voice commands and keywords of switch off function

#	Voice command	Keyword
5	The **operation** is **finished**. **Operation** is **over**. **Off** the **machines**. **Olympus /Storz off**.	▪ Operation && (finished || Off || Over) ▪ Machines && (finished || Off || Over) ▪ (Olympus ||Storz) && (finished || Off || Over) ▪ Whole system* && (finished || Off)
6	**Yes.**	▪ Yes
7	**No.** Please **cancel** closing _____(objects).	▪ No ▪ Cancel
8	Turn **off** the _____**(objects)**. **_____(objects)off**.	▪ (Endoscope || Light source ||OR table || Laser) && Off

* Added to make object keywords of switch on and switch off functions the same.

3.3 Laser lithotripter

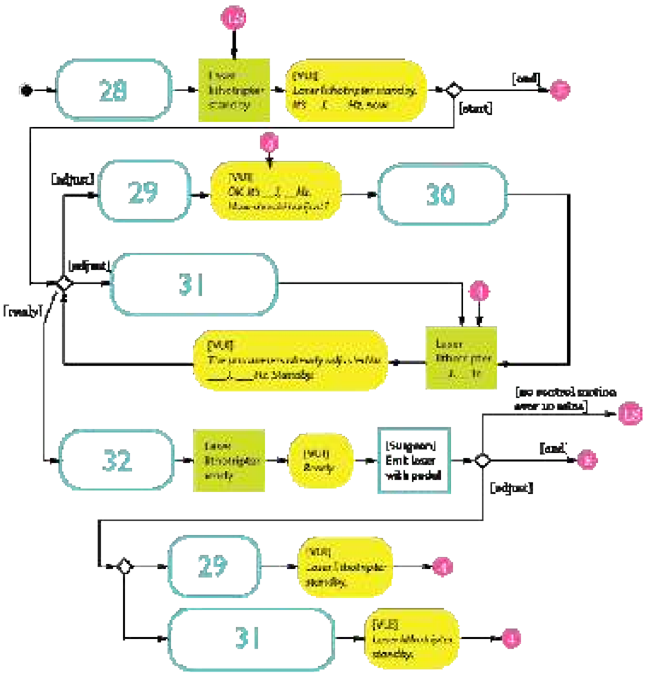

Fig. 9. The Final design of laser lithotripter

3.4 Usability Test

The usability tests were divided. In the first part, an experimental prototype was developed. The experimental prototype consisted of 3D animation and a voice recognition program. The 3D animation, which was produced using Cinema 4D, presented the first-person point of view of a surgeon during the ureteroscopy procedure. The voice recognition program was written using the Java language with help of the voice recognition open source by Google. The whole program was constructed and compiled with Eclipse software. The experimental prototype was a program that simulated a ureteroscopy scenario using Sami. Thus, the scenario of the usability test was designed to follow the scenario. Tasks, designed so that participants can use Sami's functions, fit into the scenario.

In the second part of the usability test, participants were invited to use the experimental prototype experiencing the ureterosocpy scenario with Sami. To simulate the surgery, experiment devices such as medical instruments, a laser pedal, a human model, and a Bluetooth headset were provided. In order to compare Sami with the conventional UI, the invited participants must have had surgery experience in the OR of NCKUH.

Before the experiment started, there was a brief introduction for the participant. Then, the participant sat in front of the table, with the experiment devices and a projection screen in place. During the experiment, 3D animations were projected onto a screen. The 21 frames with transparent black images and white "task" words in the animation signaled that it was time to accomplish the task at hand. Once the image was shown, the tasks were announced by the researchers. After the experiment had finished, the participant was asked to fill out a questionnaire. The participant could quantify the result using the seven-points Likert scale with three aspects: safety, cognitive load, and, satisfaction. The participant could also evaluate conventional UI and VUI by giving a score from 0 to 10.

Fig. 10. Usability testing

3.6 Comparison between VUI and conventional UI

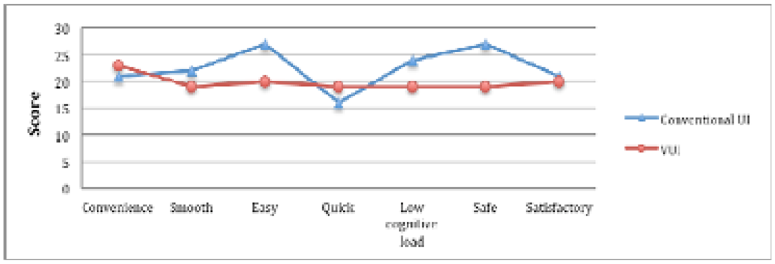

Fig. 11. The result of the comparison

Three participants evaluated two kinds of UI by scoring from 0 to 10. Scores were summed and illustrated. There were two evaluation criteria used for VUI, convenience and speed, and it received scores higher than that of conventional UI. Participants agreed that VUI could improve the efficiency of surgery, especially in a situation where medical assistants were lacking. However, other evaluation criteria of VUI, such as smoothness, simplicity, low cognitive load, safety, and satisfaction, got lower scores. The reason was that the participants encountered many machine errors -- computer crashes, computer viruses, and, disconnections -- in their daily life experiences. They indicated that sometimes medical assistants would adjust things without being asked. Even if the requests were sometimes not spoken clearly, medical assistants could still make the right decision. This might be the most important function that Sami could not achieve at that time.

4 Results

In this experiment, we invited three surgeons with 1, 5, and 29 years of OR experience, respectively, to participate. They carried out all tasks by using English voice commands to control the medical devices.

Error

Errors were recorded when a voice command given by a participant did not match the keyword. The error rate was 23.8%. It is interesting that all the participants were concerned whether the voice commands given in the WOZ test were the same as those spoken during this test. However, in fact, the voice commands spoken during the usability test did not match all the time. Besides, although the VUI framework was designed (for surgeons) to understand simple voice commands, when the function was used repeatedly, the participants did not notice this design feature.

Safety

The questions on safety listed in the questionnaire were designed according to ISO 14971. The safety score given by the participants was 63.9%. The participants did not consider that VUI caused any distractions during the surgery, and the feedback on Sami was acceptable. Besides, the accuracy of the voice recognition feature was the most important thing that concerned the participant. Even though Sami did not make any mistakes in voice recognition, over half of the participants did not trust in the voice recognition technology. This was the main reason why the safety score was so low.

Cognitive load

The cognitive load mean score of the participants was 65.8%. They felt that their interaction with Sami was straightforward and clear. Moreover, the hierarchies of VUI were simple to understand from the participants' point of view. They also thought that the designed controlling functions for adjusting the medical devices were easy to operate.

Some of the participants indicated that English was not as familiar to them as Chinese so sometimes they encountered a certain amount of cognitive load. While other participants thought that it did not bother them, VUI undoubtedly improved their concentration on the therapy. Because they did not trust the voice recognition feature, they enunciated deliberately every word they said. Though the technical issue is out of the scope of this research, it is still an influential factor affecting the scores given by the participants.

Satisfaction

The cognitive load mean score given by the participants was 75.6%. The participants were interested in the voice control functions of Sami and satisfied with the medical devices' controls. They believed that Sami could improve the experience of surgery. In addition, the experience of taking part in the experiment was excellent from the participants' point of view. However, based on their concern about surgical safety, they tended to choose the conventional UI over VUI.

5 Conclusion

Research Review

1. Consolidating the scenario and requirements of ureteroscopy in Taiwan.

Through several interviews and observations, the scenario was consolidated, and the surgeons' requirements were listed in a table. In the academic field, this design process has been implemented frequently, however, there was a dearth of studies that discussed Taiwan's medical environment.

2. Developing a new VUI of CAS system for Taiwanese medicine through new VUI design concepts.

Sami was implemented with new design concepts: multiple input, single output and natural dialogue. Moreover, there were two experiments implemented and three versions of Sami designed during the design process. Medical environments differ from country to country; so it is noteworthy that Sami was designed for Taiwanese surgeons specifically. Because the voice recognition technology for Chinese was not accurate enough, Sami was constructed using English. In the future, we would be very happy to see Chinese VUI applications of the CAS system created and implemented in every OR.

3. The usability of Sami is highlighted by implementing the usability test.

After the experimental prototype of Sami had been constructed, three surgeons were invited to participate in a usability test with the hope of experiencing a simulated OR with Sami. Through interviews and questionnaires, the participants expressed their opinions of Sami. They found that the designed elements of Sami such as audio feedback and the hierarchies were very useful. However, they lacked confidence in the accuracy of the voice recognition feature, which accounted for Sami's comparison scores being lower than those of conventional UI in most of the evaluation criteria.

6 References

1. Grange, S., Fong, T., & Baur, C. (2004). M/ORIS: a medical/operating room interaction system. Paper presented at the Proceedings of the 6th international conference on Multimodal interfaces, State College, PA, USA.
2. Irion, K., & Novak, P. (2000). Systems workplace for endoscopic surgery. Minimally Invasive Therapy & Allied Technologies, 9(3-4), pp. 193-197.
3. Kelley, J. F. (1983). An empirical methodology for writing user-friendly natural language computer applications. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, Massachusetts, USA.
4. Kortum, P. (2008). HCI beyond the GUI: design for haptic, speech, olfactory and other nontraditional interfaces: Morgan Kaufmann.
5. Perrakis, A., Hohenberger, W., & Horbach, T. (2013). Integrated operation systems and voice recognition in minimally invasive surgery: comparison of two systems. Surgical endoscopy, pp. 1-5.
6. Roe, D., & Wang, Y. (2000). A voice-controlled network for universal control of devices in the OR. Minimally Invasive Therapy & Allied Technologies, 9(3-4), pp. 185-191.