

# “I Know U” – A Proposed VUI Design for Improving User Experience in HRI

Chen Liu<sup>1</sup>, Aik Joon Sharyl Quek<sup>1</sup>, Shi Jie Alicia Sim<sup>1</sup>, and Swee Lan See<sup>2</sup>

<sup>1</sup>River Valley High School, Singapore

{liuchenqd, alicia\_simsj}@hotmail.com, sharylquek@gmail.com

<sup>2</sup>Institute for Infocomm Research (I<sup>2</sup>R), A\*STAR, Singapore

slsee@ieee.org

**Abstract.** A new aspect of the VUI design for improving user experience in human-robot interaction is proposed. By leveraging on behavioral analysis, we attempt to identify user’s ethnic group from his speech. A preliminary study on Chinese nationals, Singaporean, and Japanese showed that different ethnic group display different speech pattern. However, across ethnic groups, the spoken style of users with same personality does not differ much. Within each ethnic group, speech pattern of users with different personality can be different. We affirmed that personality influences speech and it can be determined in spoken communication.

**Keywords:** Voice User Interface, VUI, Human-Robot Interaction, HRI, behavioral analysis, personality, speech, multi-cultural communications, MySQL, PHP.

## 1 Introduction

Many people envisage that a robot can undertake challenging, or unpleasant tasks for them. However, the development in social robotics has not been able to deliver up to that expectation yet. According to Kirkpatrick and Fisher [1], the development of the robotics market closely resembles that of the personal computers market. The challenge still lies in the area of artificial intelligence and human-robot communications that would allow robots to be easily used by people.

In this paper we propose a new aspect of the Voice User Interface (VUI) design, which can be considered for improving the user experience during human robot interaction, namely to design a robot that can identify a user’s ethnic group from the user’s speech, and communicate with the user by leveraging on behavioral analysis. We present a preliminary study to discover the speech pattern of three ethnic groups, namely Chinese nationals, Singaporean, and Japanese, and study them with respect to their personality.

Multi-cultural communication is a challenging issue [2]. Programming a robot to recognize speech pattern of people of different ethnic group and personality would allow the robot to recognize and respond to the user more accurately, hence improving the efficiency and quality of human-robot communication. The claim can

be tested to develop voice animated virtual robot or personal robot for entertainment, education or even healthcare. The personal robot market in these areas has been forecasted to be a growing market of good ROI potential [1]. If the proposed design option could enhance user experience, it should improve user acceptance of social robots, and thus a significant achievement of the research project.

## **2 Method and Materials**

### **2.1 Human Subjects**

We invited students of age between 12 and 18 years old in Singapore to participate in this study. We managed to contact and engage 65 students as our human subjects. They comprises:

- 25 Japanese (7 introverts, 18 extroverts)
- 22 Chinese nationals (14 introverts, 8 extroverts)
- 8 Singaporean (11 introverts, 7 extroverts)

The participants were selected based upon their cultural backgrounds needed for this study. The Chinese nationals and Japanese participants must not have stayed in Singapore for more than 2 years, so as to ensure that they were still having their strong cultural essence from their own countries and have not been influenced by the Singaporean culture. By setting this criterion, we hope to ensure that the final result can be accurate and there was no cross cultural effects and bias.

We have both male and female human subjects, and they are identified by their cultural background and personality in this study. Their personality was determined using the DISC psychometric tool to perform human behavioral analysis. Prior to conducting the experimentation with the human subjects, approval was obtained from an institutional review board. Consent forms were given to the participants and approvals were sought from their parents/guardians.

### **2.2 Information Management**

To ease data collection, data storage and data analysis, an information system was designed and developed using PHP scripting and MySQL (see Figs. 1 and 2). The system has an administrator user interface. This controlled access to the sensitive information collected from the human subjects that are being stored in the database of this system for this study.

Fig. 1. HTML form (User interface)

Field	Type	Collation	Attributes	Null	Default	Extra	Action
<input type="checkbox"/>	online_chat_set		int(11)		No	None	
<input type="checkbox"/>	face_to_face_set		int(11)		No	None	
<input type="checkbox"/>	First_Name		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Middle_Name		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Last_Name		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Age_Group		int(11)		No	None	
<input type="checkbox"/>	Email_Address		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Contact_Number		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Gender		varchar(11)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Nationality		varchar(11)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	School		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Personality_Type		varchar(11)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Hands_forced		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Hands_by_the_side		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Others		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	continuous_eye_contact		int(100)		No	None	
<input type="checkbox"/>	broken_eye_contact		int(100)		No	None	
<input type="checkbox"/>	Fillers (type/no.)		varchar(255)	latin1_swedish_ci	No	None	
<input type="checkbox"/>	Interview set: Speaking/s		int(11)		No	None	
<input type="checkbox"/>	Interview set: Pause/s		int(11)		No	None	
<input type="checkbox"/>	Interview set: Total length/s		int(11)		No	None	
<input type="checkbox"/>	Online set: Total length/ words		int(11)		No	None	

Fig. 2. Database Structure (using PHP MyAdmin and MySQL Workbench)

### 2.3 Procedure

A field study was carried out in the form of an interview with the 65 human subjects in their schools. An invitation letter was sent to the school administrator to obtain their support to participate in this study. Upon obtaining the school’s approval, arrangement was made with the respective school teachers to conduct the experiments with the students.

Before the experimentation started, the students registered themselves through the information system created, and they were explained on the objectives and details of this research project.

In order to know the human subjects' personality types, we performed the DISC behavioral analysis on them. The DISC assessment codes were sent to the Chinese nationals and Singaporeans via email and their assessments were done online. For the Japanese participants, as reading English was a challenge for them, we translated the personality assessment into Japanese language to ease their reading and understanding. We assume there is no effect for the different cultural group to use different type of the assessment, since it was a standard psychometric tool that they were using. From this exercise, we could therefore classified the human subjects into Extrovert and Introvert users.

During the experiment, the following questions were posted to the 65 participants:

1. What do you like to do in your free time?
2. What subjects do you like? Feel free to explain why.
3. How long do you spend on the computer in a week?
4. What genre of books do you read?
5. What is your opinion on the popularity of K-pop?
6. Do you have any complaints regarding your school system?
7. Would you please describe your country's culture?
8. How do you manage you schoolwork and other commitments?
9. What do you think of school students attending school with cosmetics/make-up on?
10. If money is not an issue, what would you like to buy or do?

By answering these questions, we recorded and observed the spoken styles of each individual with respect to his or her personality and cultural background to do further analysis.

### **3 Data Analysis**

The interviews were recorded using video cameras and audio phones. The following speech patterns were guide for us to perform the data analysis:

1. Length of speech and pauses
2. Length of responses
3. The common fillers used and frequency of usage

Tables 1 and 2 below present the results after analysis.

**Table 1.** Speech patterns results of the introverts and extroverts of Chinese nationals, Singaporeans and Japanese participants

	Chinese nationals		Singaporeans		Japanese	
	<i>I</i>	<i>E</i>	<i>I</i>	<i>E</i>	<i>I</i>	<i>E</i>
Length of response/s	102	87	75	71	29	29
Length of pause/s	35	26	34	18	22	22
Length of speech/s	67	61	41	53	7	7
Percentage of time spent on speech/ %	66	70	55	75	59	59
No. of fillers used	8	12	8	8	1	2
<i>I: Introvert ; E: Extroverts</i>						

**Table 2.** Analysis on fillers from the experiment

Speech Pattern	Chinese nationals	Singaporean	Japanese
Length of answer/s	long	medium	short
Top 3 fillers used	1	'Like'	'err...'
	2	'yeah'	'oh'
	3	'ah'	'like'
			'eetto'
			'ma'
			'ssii'

## 4 Results and Discussion

It was found that different human subjects exhibit different speech patterns. Participants of the same cultural group tend to display similar speech patterns despite personality difference.

For Chinese nationals, if the length of response, length of pause and length of speech is generally long, yet if the number of fillers used is generally few, s/he may be a Chinese introvert. To further confirm the personality, Chinese introvert should have more discontinuous eye contact than continuous eye contact. We also found that the Chinese can be identified as an introvert if s/he has fewer tendencies to fold hands, wave and touch hair. However, if the length of response, length of pause and length of speech is generally short, yet if the number of fillers used is generally more, he or she may be a Chinese extrovert. To further confirm the personality, Chinese

extrovert should have more continuous eye contact than discontinuous eye contact. We can also further identify the Chinese as an extrovert if he or she has a greater tendency to fold hands, wave and touch hair.

For Singaporeans, if the length of response and length of pause is generally long, but the length of speech is moderate, he or she may be a Singaporean introvert. To further confirm the personality, Singaporean introvert should have more discontinuous eye contact than continuous eye contact during the speech. We can also further identify the Singaporean as an introvert if he or she has fewer tendencies to fold hands, wave but a greater tendency to clench hands together. However, if the length of response and length of pause is short, but the length of speech is long, he or she is potentially a Singaporean extrovert. To further confirm the personality, Singaporean extrovert should have more continuous eye contact than discontinuous eye contact during the speech.. We can also further identify the Singaporean as an extrovert if he or she has a greater tendency to fold hands, wave but fewer tendencies to clench hands together.

Within each ethnic group, we have also observed that the speech pattern (i.e. speech and pauses) of the introverts and extroverts can be different. However, across ethnic groups, there is little difference in terms of the speech patterns of the speakers of same personality type. We therefore affirmed that personality influences speech [3] and can be determined in spoken communication [4]. Language is a vehicle of personality and effective speech goes hand in hand [5].

We would therefore like to propose a VUI design that makes a robot appear to be a smart pal to consumers. The underlying principle is that the robot knows the user (see Fig. 3), and the user can establish a close and trust-worthy relationship with the robot. The robot is able to identify speaker’s ethnic group and personality, and the robot could be more competent and efficient in communicating with the user, allowing the user to have a better user experience through the interaction.



Fig. 3. Proposed “I Know U” VUI Design to Improve User Experience in HRI

## 5 Future Directions

A new aspect of the VUI design for improving user experience in human-robot interaction has been proposed in Fig. 3 above. In conclusion, people of different

culture and personality displayed different speech patterns and behaviors, such as body movements and eye contact, when communicating. We hope to continue this study by investigating into more aspects of human communication through interviews. We could also research more deeply into the language and communication aspects in different Asian cultures and interview more participants from different Asian cultures. Further studies can be carried out to investigate the other factors affecting speech patterns besides personality and culture, such as gender, age, etc. and we could use more elaborated analytical methods. Last but not least, we hope the proposed VUI could be put into place, and more insights can be captured from the human-robot interaction to increase the productivity and user-friendliness of the personal robot market.

## References

1. Kirkpatrick, K., Fisher, L.: Personal Robotics: Task, Security and Surveillance/Telepresence, Entertainment and Education Robots, and Robotic Components Markets Through 2017. ABI Research (2010)
2. Markel, P.: Cultural Differences Impact Workplace Communication (2009), [http://www.alliancetac.com/?PAGE\\_ID=2560](http://www.alliancetac.com/?PAGE_ID=2560)
3. Polzehl, T., Möller, S., Metze, F.: Automatically Assessing Personality from Speech. In: IEEE Fourth International Conference on Semantic Computing, pp. 134–140 (September 2010)
4. Friedrich, P., Redfield, J.: Speech as a Personality Symbol: The Case of Achilles, Linguistic Society of America, University of Chicago (1978)
5. Metallinou, A., Lee, C.C., Kazemzadeh, A., Busso, C., Lee, S., Narayanan, S.: Interpreting ambiguous emotional expressions (September 2009)