User Experience Factor Investigation of a Voice User Interface (VUI) for the Elderly

Adriance Wilfred
Seoul National University
adriance@snu.ac.kr

Minjee Kim
Seoul National University
mj.kim1210@gmail.com

Sungho Kim
Seoul National University
ksh32035@snu.ac.kr

Chan Hyeok Yun
Seoul National University
ych8184@snu.ac.kr

NamWoo Cho
Seoul National University
chonamwoo@snu.ac.kr

Jane Lee
Seoul National University
lee.jane@snu.ac.kr

Soo Yeon Kim
Seoul National University
sykim213@snu.ac.kr

Joong Hee Lee
Seoul National University
jetishie@snu.ac.kr

Kyung-Jun Lee
Seoul National University
seerjun@snu.ac.kr

Myung Hwan Yun
Seoul National University
mhy@snu.ac.kr

Abstract

This paper presents an investigation of UX factor of a voice user interface (VUI) for the elderly, where the voice user interface in this study refers to the artificial intelligence (AI) based technology that interacts with a user via speech commands. In this research, the investigation of UX factors involves the acceptance and affective aspects of the elderly. We applied the Wizard-of-Oz methodology in this experiment, and established our result using Exploratory Factor Analysis (EFA) which determines the significant elements of user experience in voice user interface. Overall, the extracted UX factors can be regard as relevant design parameters for future voice user interface system and researches.

Keyword

Voice interaction, User experience, Artificial intelligence, Elderly people

1. Introduction

1.1 Background

It is inevitable that many elderly people are facing issues such as social isolation and loneliness. Elderly who suffers from social isolation and loneliness may be imposed to decrement of their cognitive mental state as their physical ability decreases along side by side (Shibata & Wada, 2011). Due to the social demands to take care of the elderly and psychological demand for independence among the elderly, voice user interface is expected to play a vital role in helping the elderly to suppress their loneliness.

Voice user interface is a speech agent with spoken dialogue systems, dialogue management systems, and embodied conversation agents. In this study, we dedicated the voice user interface as an artificial intelligence (AI) technology that lets the users to interact via speech commands. However, the use of voice user interface among the elderly had shown that the elderly had major concerns such as anxiety about machines and maladjustments (Ganyo, Dunn & Hope, 2011; McLean, 2011; Wu et al.,2014). As a matter of fact, voice user interface contributes many opportunities for the elderly to address their cognitive and emotional difficulties.

Voice user interface were commonly measured and investigated by the two major UX factors which are acceptance and affective factors. Acceptance measures utilitarian and hedonic factors to determine the overall acceptance of the voice user interface system (De Graaf & Allouch, 2013), whereas affective factors were measured using the Positive and Negative Affect Schedule (PANAS) (Watson, Clark & Tellegen, 1988).

1.2 Acceptance factor and affective factor

Understanding the reasons for rejecting or accepting a conversational speech system is important to improve system designs and implement elaborate strategies (Wu et al., 2014). It is necessary to apprehend the influence of acceptance factors in

order to narrow down the great gap created between available technologies and the anxiety towards machines among the elderly.

Apart from acceptance aspects, the affective aspects have quite the equivalent significance in emphasizing the design of voice interfaces in human-robot interaction according to recent studies (Chanel, Kierkels, Soleymani, & Pun, 2009). Providing companionship using voice user interface in improving the quality of life for elderly and may lessen social isolation and loneliness issues among the elderly (Broekens, Heerink, & Rosendal, 2009).

This study aimed to investigate the differences in acceptance factors and reflect its result to suggest ways to increase elderly's usability when using voice user interface, and investigation of elderly's emotion state through voice interaction with the voice user interface is required because the future use of voice interfaces must be established based on the affective factors when using speech system

2. Method

2.1 Subjects

There were 53 subjects (25 male and 28 female), ranging the age from 60 to 82 years old with an average of 69.14 years of age participated in the experiment. All of the subjects have no experience with voice user interface by qualifying only those who without prior experience. Among 53 subjects, 2 of the subjects were exempted from the final analyses and their experimental sessions were halted due to their inability to hear and perceive the voice of the device.

2.2 Dependent variables

In this study, the acceptance and affection of voice user interface acceptance are measured using a questionnaire of items. For acceptance factors (Table 1), 11 factors in questionnaire prepared that require the subjects to rate on a seven-point Likert scale (1 = strongly disagree to 7 = strongly agree), whereas affective factors (Table 2) had 20 factors in the questionnaires to be rated on a five-point Likert scale (1 = strongly disagree to 7 = strongly agree).

Table 1. Acceptance factors and their questionnaire items

Factor	Definition	Questionnaires
Anxiety (ANX)	Anxious or emotional reactions evoked when using the system	If I used the conversational robot, I would be afraid to make mistakes with it
		If I used the conversation robot, I would be afraid to talk with the robot
Attitude toward using the conversation robot (ATTR)	Feelings about applying the technology	I think it's a good idea to use the conversation robot.
		It's good to make use of the conversation robot.
Intention to use (ITU)	The intention to use the system over time	I think I'll use the conversation robot during the next few days.
		I will certainly use the conversational robot again.
Perceived adaptability (PAD)	The perceived ability of the system to adapt to the users' needs	I think the conversation robot can be adapted to my needs.
		I think the conversation robot will help me when I consider it necessary.
Perceived enjoyment (PENJ)	Feelings of pleasure associated with using the system	I enjoy when the conversation robot talks to me.
		I find the conversation robot enjoyable.
Perceived ease of use (PEOU)	The degree to which one believes that using the system is free of effort	I think I can use the conversation robot without any help.
		I find the conversation robot easy to use.
Perceived sociability (PS)	The perceived ability of the system to engage in sociable behavior	I find the conversation robot pleasant to interact with.
		I think the conversation robot is nice.
Perceived usefulness (PU)	The degree to which a person believes that the system would be assistive	I think the conversation robot is useful for me.
		It would be convenient for me to have the conversation robot.
Social	The perceptions of	I think many people

influence (SI)	significant people regarding use of the system	would be pleased if I had the conversation robot.
		I would like to show my use of the conversation robot to many people around me.
Social presence (SP)	The experience of sensing a social entity when interacting with the system	When interacting with the conversation robot, I felt like I was talking to a real person.
		Sometimes the conversation robot seems to have real feelings.
Trust (TR)	The belief that the system performs with personal integrity and reliability	I would trust the conversation robot if it gave me advice.
		I would follow the advice the conversation robot gives me.

Table 2. Affective Factors

Interested	Irritable	Distressed	Alert	Excited
Ashamed	Upset	Inspired	Strong	Nervous
Guilty	Determined	Scared	Attentive	Hostile
Enthusiastic	Jittery	Active	Proud	Afraid

2.3 Experimental Environment

This experiment implemented the Wizard-of-Oz research method, where the subjects would believe the voice user interface fully able to have conversations with the them (Porter et al., 2013). A partition was set up as a divider, and placed the experimenter and subjects at each opposite side of the facing the partition (Figure 1). AI speaker that has the voice user interface function was placed on the subject's table, and a computer was placed at the experimenter's table so that the experimenter could send the pre-recorded voice files to the AI speaker as an output to allow smooth and natural interaction. The partition concealed the experimenter at the other side as the experimenter gave manual control without the subjects being aware.

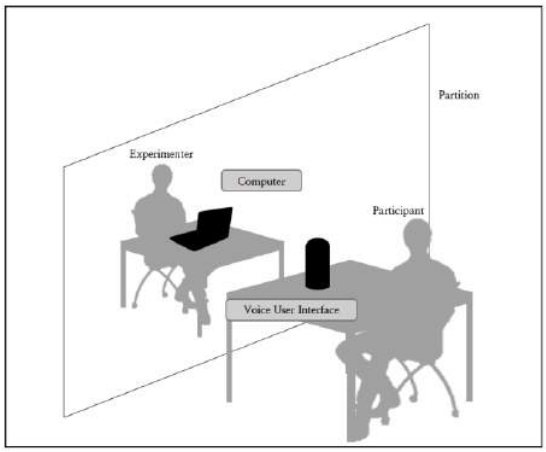

Figure 1. Experimental set up for the Wizard-of-Oz method

2.4 Procedure

The subjects practiced interacting with the voice user interface using simple randomly assigned conversational questions, such as "How is the weather?" and "What day is it?". The experimenter only entered the subject side of the partition to provide explanation of the experimental procedures, and left the subject's side as if the subject was alone with the speaker. The topics were randomly assigned using a random number table. The subjects asked to facilitate about eight conversations, with an interval of two to three minutes of break time during the experiment. In case if there were any inevitable erroneous situations, the experimenter will interrupt the experiment and halt if necessary. The halted experiment's data were then to be excluded from the final analyses.

The subjects completed a set of questionnaires to evaluate their acceptance and affective state towards the voice user interface. Overall, the experiment took approximately 70 minutes.

3. Results.

3.1 Factor Analysis of UX Elements

Based on the two UX factors aforementioned in this study, both the acceptance factor (11 elements) and affective factor (20 elements) were reduced its dimension using Exploratory Factor Analysis (EFA). Components were extracted through the exploration of the factor structure.

3.2 Acceptance factor analysis

The 11 elements of acceptance were analyzed using exploratory factor analysis. As a result, all of the factors were categorized into two components (Table 3.2) with eigenvalues greater than 1, accounting for 70.94% of the variance, each factor loading was between 0.592 and 0.874.

From the result, the first component extracted were include social presence, perceived sociability, trust, social influence, and perceived usefulness. The second component included perceived ease of use, perceived adaptability, anxiety, perceived enjoyment, intention to use, and attitude. There were noticeable certain features of the elements included on each component. All of the composed first component elements were viewed as "intimacy" factors which were related to social factor between the user and the voice user interface. The elements in second component were assessed as "user engagement" factors which can be determined as the evaluation of the system's usability values.

Table 3. Results of factor analysis for acceptance factors

Components	Items	Factor loadings	Eigenvalue	Extracted variance
Intimacy	SP	.855	6.508	59.167
	PS	.800		
	TR	.791		
	SI	.759		
	PU	.657		
User engagement	PEOU	.874	1.295	11.773
	PAD	.773		
	ANX	.755		
	PENJ	.674		
	ITU	.631		
	ATTR	.592		

3.3 Affective factor analysis

Similarly, with the acceptance factor, the 20 elements of affective were analyzed using exploratory factor analysis. The result was classified into five components (Table 4).

The five components extracted with eigenvalues greater than 1, accounting for 71.25%, and each factor loading was between 0.555 and 0.846. The first component included guilty, distressed, nervous, ashamed, afraid, upset, and scared. The second component included jittery, alert, irritable, and hostile. The third component included enthusiastic, interested, active, and proud. The fourth component included strong, inspired, and excited. The fifth component including attentive and determined. There were noticeable traits of the factors included on the components. All of the composed first component factors were viewed as "internal-negative" factors which associated with the internal negative emotions, whereas composed second component factors were clustered as "external-negative" factors which the external negative emotions towards the system such as jittery and irritated. Composed third component were identified as "external-positive" factors which positive emotions were externally, and composed fourth component are the "internal-positive" factors where users having internal positive emotions such as being inspired and excited. The last component was defined as "awareness" factor as the attentive and determined factors can be identified as the awareness that the voice user interface system exists, and the users are able to initiate conversation with the voice user interface system.

Table 4. Results of factor analysis for affective factors

Components	Items	Factor loadings	Eigenvalue	Extracted variance
Internal-negative	Guilty	.844	5.967	29.837%
	Distressed	.790		
	Nervous	.769		
	Ashamed	.739		
	Afraid	.626		
	Upset	.598		
	Scared	.573		
External-negative	Jittery	.792	3.579	17.895%
	Alert	.790		
	Irritable	.729		
	Hostile	.555		

External–positive	Enthusiastic	.803	2.116	10.581%
	Interested	.773		
	Active	.721		
	Proud	.676		
Internal–positive	Strong	.846	1.412	7.060%
	Inspired	.735		
	Excited	.604		
Awareness	Attentive	.728	1.178	5.890%
	Determined	.667		

4. Discussion

Based on the result, two of the major UX factors of voice user interface for elderlies were investigated; questionnaire items of both the acceptance factors and affective factors were collected and reduced in their dimensions via Explanatory Factor Analysis (EFA). There were two components for the acceptance factor: intimacy, and user engagement. On the other hand, there were five components for the affective factor: internal–negative, external–negative, external–positive, internal–positive, and awareness (Figure 2).

The intimacy component is neither the utilitarian nor the hedonic quality; it may be a new acceptance factor to consider for the elderly users, especially. It is highly linked to the relationship–based interaction between the user and the voice user interface. Social presence (SP) holds the strongest association to the underlying intimacy component by having the highest factor loading of 0.855. Elderlies require social relationship, in other word rapport, to communicate with the voice AI agent as if they would with human agent.

The user engagement component, however, can be associated with the usability of the voice user interface system by elderly. This component consists of both the utilitarian and hedonic quality (De Graaf & Allouch, 2013). Perceived ease of use (PEOU) strongly correlates with the user engagement component, which valued the highest factor loading of 0.874. Since they are not familiar with the system, the ease of use plays the most essential role for them to engage with the system. One of the elements in the user engagement component is “anxiety”. It indicates the behavior of elderly user toward the new technology as the result confirms with the previous studies regarding anxiety (Ganyo, Dunn & Hope, 2011; McLean, 2011; Wu et al.,2014). There may be a psychological barrier against the voice user interface for the elderly to engage with the system since it is unfamiliar and unadaptable.

The affective factors can be widely categorized into two components; negative and positive components, and these components were then specified into whether they are internal and external. In terms of negative components, guilty holds the highest factor loading (0.844) for internal–negative, whereas jittery (0.729) for external–negative. The internal–negative component indicates their negative emotional states toward themselves when they interact with the system. They can feel guilty, distressed, ashamed, or upset. The elderly users rather tend to blame themselves when the error occur. At the same time, in terms of external negative component, the user shows their negative emotional states against the system when there is an error occurred. On the other hand, for positive components, strong holds the highest factor loading (0.846) for internal–positive, and lastly enthusiastic (0.803) for external–positive. User may be internally excited to interact with a new system in positive hand. Moreover, externally, the user can also be emotionally engaged with the system to show feel their enthusiasm and interest as they interact with the voice user interface system. Interestingly, the affective component did not rely on the physical design of the voice user interface system, but with the interaction between the elderly user and the system. Therefore, it is more important to deliver positive user experience by providing a good interaction and feedback within the voice user interface system. Such result is probably because of the characteristic of the voice user interface system that the users do not physically interact with the system but with their verbal communication only so the physical appearance is not the determining factor for the affective components.

Concurrently, the awareness component is categorized as neither positive or negative. This component consists of attentive and determined elements. Attentive placed the highest factor loading of 0.728 which greatly suggests that it holds strong correspondence to awareness which can be interpreted as the voice user interface is always there for the user for conversation and able to give feedback. Besides this indicates that the elderly user must recognize that there is a system they can communicate with, otherwise there will be no interaction made. Unlike the social presence, its physical presence matter in this case; the system must let the elderly user know it exists.

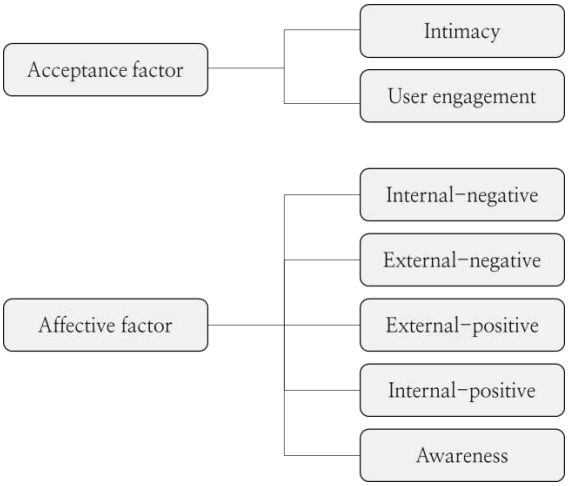

Figure 1. UX factors for voice user interface by EFA

5. Conclusion

Based on these factors investigated, one can propose the design of a future voice user interface system for the future research. The future study should focus on the design parameters relevant to the extracted UX factors: intimacy, user engagement, internal and external emotional state (both negative and positive), and the awareness. Moreover, when considering that the users are elderlies, there must be a new design to overcome the psychological barriers within the users to accept the new system.

Overall, this study provides insights with regards to acceptance and affective factors and to help the manufacturers or designers to design the voice user interface system for the elderly to improve their quality of life against social isolation and loneliness.

Reference

1. Broekens, J., Heerink, M., & Rosendal, H. (2009). Assistive social robots in elderly care: a review. Gerontechnology, 8(2), 94-103.
2. Chanel, G., Kierkels, J. J., Soleymani, M., & Pun, T. (2009). Short-term emotion assessment in a recall paradigm. International Journal of Human-Computer Studies, 67(8), 607-627.
3. De Graaf, M. M., & Allouch, S. B. (2013). Exploring influencing variables for the acceptance of social robots. Robotics and Autonomous Systems, 61(12), 1476-1486.
4. Ganyo, M., Dunn, M., & Hope, T. (2011). Ethical issues in the use of fall detectors. Ageing & Society, 31(8), 1350-1367.
5. McLean, A. (2011). Ethical frontiers of ICT and older users: cultural, pragmatic and ethical issues. Ethics and information technology, 13(4), 313-326.
6. Shibata, T., & Wada, K. (2011). Robot therapy: a new approach for mental healthcare of the elderly - a mini-review. Gerontology, 57(4), 378-386.
7. Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of personality and social psychology, 54(6), 1063.
8. Wu, Y. H., Wrobel, J., Cornuet, M., Kerhervé, H., Damnée, S., & Rigaud, A. S. (2014). Acceptance of an assistive robot in older adults: a mixed-method study of human-robot interaction over a 1-month period in the Living Lab setting. Clinical interventions in aging, 9, 801.