Intelligent Social Robots in the Wild:

A Qualitative Study on Deploying Intelligent Social Robots with Growing Features to Real Home Environments

Chan Mi Park, Yuin Jeong, Kwangmin Jeong, Hae-Sung Lee, Jeehang Lee, Jinwoo Kim, *Member of ACM*

Abstract— **With an ambience of artificial intelligence and robot technologies, humans are surrounded by various social robots. This work-in-progress explores sound recognition robot that is currently being developed to be sold to single family households in 2018. Before public launch, nine units had been built and deployed into nine houses for two weeks. We would like to compare two types of intelligence: Ambient Intelligence (AmI) and Attentive Intelligence (AtI). With a growth metaphor adapted, nine field participants have experienced in teaching the intelligence of the robot for one-week (AtI state), and using the fully-grown robot (AmI state) on the second week. As a result, despite having many annoying and negative experiences, participants have chosen to have the robot with AtI.**

I. Introduction

Enhancing human-robot interaction more natural and sociable have been a prolong issue in various studies in robotic and HCI community [8]. With an advanced technology of artificial intelligence (AI), the presence of AI is becoming more transparent and pervasive in social robots. Ambient Intelligence (AmI) best describes current level of system. Philips Research describes AmI as "a digital environment that supports people in their daily lives in a nonintrusive way". Its highlighted features, sensitive, responsive, adaptive, transparent, ubiquitous, and intelligent, are applied in context-aware computing [7]. However, we want to raise a question about highly transparent and sophisticated AI being an optimal solution for a social robot used in uncontrolled human environment. This research explores user's perception on Attentive Intelligence (AtI) and Ambient Intelligence (AmI) by utilizing social robot, Fribo in the real home environments for two weeks before public launch.

II. Home Social robot, Fribo

A. Overview

This research utilized the social robot, "Fribo", which firstly identifies user's home activities through sound recognition from the living noise data and delivers the activity information to other Fribo users afterwards. The primary function of "Fribo" is to share the daily activities to alleviate isolated feeling and loneliness that one experiences while living at home alone (see Jeong *et al.* (2018) for more details).

Fribo is developed on Raspberry Pi 3 and its configuration is divided into input devices (microphone, ultrasonic sensor, a sound sensor, an illuminance sensor, temperature and humidity sensor) and output devices (LCD and two speakers) shown in Figure1.

Given the synthesized input sensors, Fribo operates a 'living noise recognizer' in order to detects several sounds of home appliances e.g. vacuum cleaner motor running, microwave, washing machine, refrigerator door, entrance door and room door etc. 'Living noise recognizer' is a sound data classifier using a machine-learning algorithm. With the collected acoustic data for each activity, we trained the living noise recognizer using Support Vector Machine (SVM).

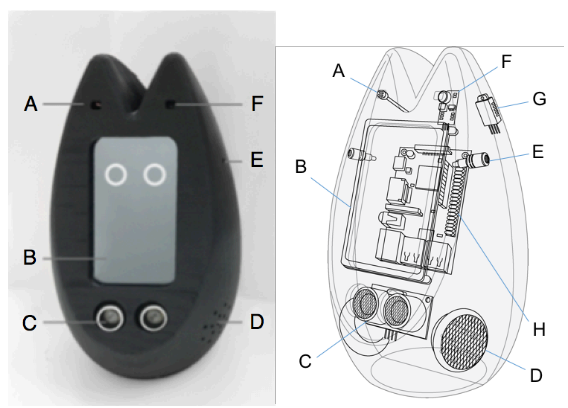

Figure 1: Fribo consists of (A) Illuminance sensor, (B) LCD, (C) Ultrasonic sensor, (D) Speaker, (E) Microphone, (F) Sound sensor, (G) Temperature and humidity sensor, (H) Raspberry Pi 3.

B. Growth interaction design of Fribo

It is important to determine what aspects (e.g. speech, emotion appraisal and exhibition, intelligence, etc) of robot users perceive as growth. In Fribo's case, the intelligence delivered in the form of speech and visually exhibited emotion could be seen as growth criteria. According to King's study (1996), a mismatch of agent's actual performance and its perceived intelligence can cause negative user experience [10]. Thus, it is necessary to provide user a

C. M. Park is with Yonsei University, Seoul Korea parkmimymo@gmail.com, Yuin Jeong is with Yonsei University, Seoul Korea, youin_v@naver.com, K. M. Jeong is with Yonsei University, Seoul Korea, vgb@gmail.com, H.S. Lee is with Yonsei University, Seoul, Korea geneel@me.com, J.H. Lee is with KAIST, Korea jeehang@kaist.ac.kr, Jinwoo Kim is with Yonsei University, Korea (corresponding author +8210-6307-2528, jinwoo@yonsei.ac.kr)

	Speech						Facial expression	
	Speech rate	Filler phrase	Speech error	Pitch	Context length	Personality	Static / Dynamic	Emotional expression
Lv. 1	Low	X	O	Low	Short	X	Static	X
Lv. 2	High	X	X	Low	Short	X	Dynamic	Low
Lv. 3	High	O	X	High	Long	X	Dynamic	Low
Lv. 4	High	X	X	High	Long	O	Dynamic	High

Table 1: Speech parameter and facial expression adjustment for different growth levels

robot with low perceived intelligence in the beginning of usage, but manifests its advanced intelligences as it progressively evolves.

Fribo's perceived intelligence is classified into four levels of growth levels, each level composed of synchronized visual and auditory elements. As the perceived intelligence and the proficiency [4] of a robot can be controlled by manipulating the speech fluency [1, 2, 3], we adjusted the Fribo's speech elements including its facial expression delivering emotion exhibition and activity information (Table 1). Figure 2 represents the result of manipulation check that was conducted to measure perceived intelligence for each growth level. A total of 25 participants (male =4, female = 21) of manipulation check identified differences between each growth level and evaluated the perceived intelligence as progressive. One of unique differences is adapting human-like feature; filler phrases, such as "ah, um" were placed in-between phrases for the level three. On level four, filler phrases were excluded to deliver a fluent and clean speech to the user.

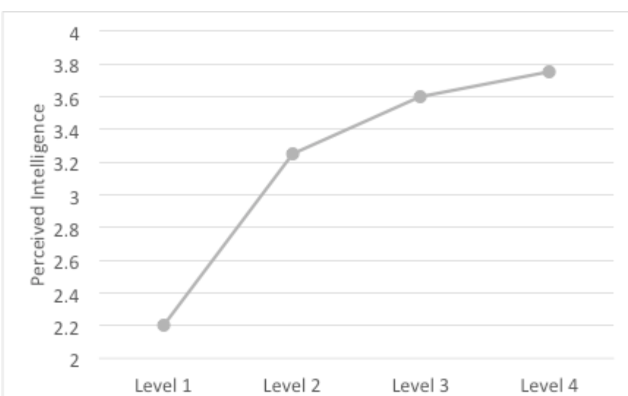

Figure 2: Perceived intelligence for each growth level.

Fribo is composed of four different facial expressions: blinking (default), like, curious, sad. According to Cassell (1994), to deliver a clear communication, it is essential to have non-verbal elements, such as facial expression, eye gaze, other bodily gesture, and pause synchronize with a speech context [6]. In order to provide more active and realistic interaction to the user, we adjusted Fribo facial expressions to emphasize the concept of growth. Level 1 is consisted of eye blinking; level 2 and 3 has blinking and smiling; and level 4 with all expressions. The growth metaphor on the robot interaction design was implemented as the key element for users to intuitively perceive the robot as AtI; and the fully-grown state (Lv. 4) as AmI.

We utilized separate smartphone application as a main channel for users to teach Fribo. When Fribo detects the event sound, a notification asking about specific event category is sent to the user as of speaking about the event detection. Fribo uses the push message to provide a question (e.g. "Which one of these sounds have I just heard?") and options for its detected event sound. For every push message, there are three choices that the user can choose, "Yes", "No", or "I don't know". If one chooses "Yes", the sound data is labelled and stored on the server. On the other hands, if the detected sound has not appeared, and user chooses "No" for the answer, Fribo then provides another question with multiple choices of all other events. As the user diligently answers questions by Fribo, the user receives one growth credit and Fribo is gradually advancing to the next level. Conversely, not answering Fribo's message after 10-minute, Fribo's growth credit is detracted and slowly degenerated to the pre-stage.

III. PROCEDURES

For the field study, we recruited three groups of three (n=9) participants via several online communities of colleges and universities in Seoul, South Korea. Recruited participants are females and their average age is 23.7 years. We used Emotional intimacy scale [11] in a survey format to measure their intimacy to each other and confirmed that their intimacy was high.

Before starting the field study, participants were asked to record the sound data generated within their home. Nine activities (refrigerator, front door, computer, window, lighting, stovetop, microwave, vacuum cleaner, washing machine, and tap water) were collected in order to train Fribo's living noise recognizer. This was practiced to compensate the limitation of Fribo's existing recognition accuracy in the real environments. Next, a week before the start of field study, Fribo was installed at each participant's house. Then, detailed guidelines on the robot and field study procedure were distributed. Participants were asked to complete a preliminary questionnaire, which consists of demographic information, privacy concerns, loneliness levels, intimacy with friends, and experiences with raising pets or plants. Until the start of field study, any issues relating to the robot, such as sensor malfunctions, and Internet connectivity issue were amended.

The field study was conducted for two weeks using within subject design. On the first week, Fribo with AtI was distributed, and on the second week after the full growth (Lv. 4), participants had an experience with Fribo with AmI. There were no specific instructions in relation to Fribo usage given to participants. During the two-week experiment, participants were interviewed twice (first interview after one week, second interview after the completion of the experiment) and conducted online surveys once every three days. Upon completion of the field study, each participant was paid a participation fee of US $230.

IV. Results

The interview from the first and second week of the field study asked users about social robots, growth and degenerated interaction, human-aid (providing feedback for the robot's question), and error recognition. The collected interview data went through a clustering process to discover the insights with respect to user's perceived intelligence of Fribo with AtI (first week) and AmI (second week). Despite instability and inaccuracy of Fribo with AtI on the first week, users commented the AtI as better experience than that in the second week with AmI robot. Following interview data are the key comments that reflect the general user preference for the AtI robot

A. personification

"*It seems like giving 'I don't know' feedback may hurt its feeling, so I did not provide that answer*" (P1, week 1)

"*When the Fribo said, 'It would be nice to have something to nibble', the robot seemed to have self-awareness*" (P1, week 2)

B. Presence

"*Because of Fribo, I felt someone's presence in my house. So when I came back home, I say 'Hey, I am home' out loud for it to hear."* (P4, week 1)

"*Since there is very little push message asking questions compared to the first week, I almost forgot about its existence. Right now, it feels just like any other home appliance*" (P3, week 2).

C. Responsibility and Ownership

"*I felt responsible for giving the right feedback.*" (P8, week 1).

"*I did not want 'my baby (Fribo)' to fall behind of other friends' Fribos. So I felt some sort of responsibility to raise the robot properly.*" (P3, week 2)

D. Emotional Attachment

"*The robot was so cute that I named it with 'Blinker'*". (P6, week 1)

"*I feel some intimacy with the robot. If I imagine the robot not being on my desk one week after, I'd feel some emptiness.*" (P3, week 1)

D. Perception on error

" *In the beginning of usage, I understood its immaturity and accepted any errors that it makes. Because it just started to learn!*" (P4, week 1)

"Because there wasn't much to give feedback about, I felt I was neglecting. I wanted to have some form of interaction with the robot, but I couldn't." (P3, week 2)

While raising and teaching the robot in the first week (i.e. AtI stage), users created attachment, responsibility, and ownership. This, in effect, made users perceive the robot's error with more understanding and forgiveness. Whereas, on the second week (i.e. AmI stage), the robot has become fully grown and turned its intelligence to an ambient one, but has not formed as high intimate relationship with the user as in the first week. In the worst case, some participants treated the robot as any other home appliance. Field participants described the robot as inactive and shown less interest and weak attachment for the robot. Even the sense of responsibility for the robot's growth has stopped during the usage of AmI robot.

V. Conclusions and Discussions

As of concluding the field study result, there is a definite need for a constantly evolving robot with AtI; Personally guided machine-learning of the AI allows the robot to learn about its living environment, and able to draw user's close attention and build much closer 'sociable' human-robot interaction. As the interview data manifests, users have come to treat the robot as a living creature while providing human-aid (feedback). Annoyance, overwhelmingness, and heavy sense of obligation were first faced as negative experience, but as the users were exposed to fully-grown state of the taciturn robot (AmI), those past negative experiences were missed.

Even though there are several valuable insights discovered from the research, there are still several shortcomings. Frist, as the field study has utilized the robot with specific function, e.g. living noise recognition, it has some limitations in delivering unique aspects of AtI and AmI. Second, the implemented growth levels are completely dependent on user's active interaction with Fribo, thus, it was difficult to define the exact timeline of when AmI has been put into action. Third, the field study is comprised with limited number of participants. As the group is focused on females in 20s, the wide range of users' insights was not investigated. Lastly, the field study on the robot with AtI and AmI was conducted with within subject. Therefore, firstly implemented AtI robot may have some influence of novelty effect to users.

For the future research the questions of *how deep should the social robot's attentive intelligence grow and last,* and the robot's performance on its degeneracy will be explored, which we believe will be the important factor for user experience with the intelligent social robot.

Acknowledgment

This material is based upon work supported by Basic Science Research Program through the National Research

Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1B02015987).

REFERENCES

[1] Martha W. Alibali, Dana C. Heath, and Heather J. Myers. 2001. Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. Journal of Memory and Language 44, no. 2: 169-188.

[2] Sean Andrist, Bilge Mutlu, and Adriana Tapus. 2015. Look like me: matching robot personality via gaze to increase motivation. In Proceedings of the 33rd annual ACM conference on human factors in computing systems, pp. 3603-3612.

[3] Kazumi Aoyama, and Hideki Shimomura. 2005. Real world speech interaction with a humanoid robot on a layered robot behavior control architecture. In Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on, pp. 3814-3819.

[4] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. International journal of social robotics 1, no. 1: 71-81.

[5] Justine Cassell, Catherine Pelachaud, Norman Badler, Mark Steedman, Brett Achorn, Tripp Becket, Brett Douville, Scott Prevost, and Matthew Stone. 1994. Animated conversation: rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents. In Proceedings of the 21st annual conference on Computer graphics and interactive techniques, pp. 413-420. ACM.

[6] Diane J. Cook, Juan C. Augusto, and Vikramaditya R. Jakkula. 2009. Ambient intelligence: Technologies, applications, and opportunities. Pervasive and Mobile Computing 5, no. 4: 277-298.

[7] Jerry Alan Fails, and Dan R. Olsen Jr. 2003. Interactive machine learning." In Proceedings of the 8th international conference on Intelligent user interfaces, pp. 39-45. ACM.

[8] Kwangmin Jeong, Jihyun Sung, Haesung Lee, Aram Kim, Hyemi Kim, Chanmi Park, Youin Jeong, JeeHang Lee, and Jinwoo Kim. 2018. Fribo: A social networking robot for increasing social connectedness through sharing daily home activities from living noise data. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (in press).

[9] William Joseph King, and Jun Ohya. 1996. The representation of agents: Anthropomorphism, agency, and intelligence. In Conference Companion on Human Factors in Computing Systems, pp. 289-290. ACM.

[10] Vaughn G. Sinclair, and Sharon W. Dowdy. 2005. Development and validation of the Emotional Intimacy Scale. Journal of Nursing Measurement 13, no. 3: 193.

[11] Andrea L. Thomaz, and Cynthia Breazeal. 2006. Transparency and socially guided machine learning. In 5th Intl. Conf. on Development and Learning (ICDL).