Investigating the Effects of Robot Voice Pitch, Robot Voice Gender, and User Gender on User Perception of Teleoperated Robots

Sichao Song[12], Jun Baba[12], Junya Nakanishi[2], Yuichiro Yoshikawa[2], Hiroshi Ishiguro[2]

[1]CyberAgent, Inc.
[2]Osaka University

Abstract: It has been suggested that robots speaking in a high pitch could be perceived as more attractive than those speaking in a low pitch. However, it is not clear whether the use of a high-pitched voice favors teleoperated robots as well. To investigate this aspect, we conducted an experiment to study the effects of the robot voice pitch, robot voice gender, and user gender on the attitudinal responses of the users toward a teleoperated robot and the associated decision-making. It was observed that the male and female participants perceived a high-pitched voice differently. The users' awareness of the robot being teleoperated and persuasiveness of the robot were noted to be related, which may provide a plausible explanation for the interaction effects between the voice pitch and user gender.

Introduction

Communication robots are playing an increasingly significant role in the present society. In Japan, owing to the problems of aging and depopulation, communication robots are being introduced to provide services in places such as airports [17], shopping malls [7], and hotels [12], which often attracts customers. Because the task of a communication robot may not include only chitchatting but also providing product recommendations to people, it is important for such a robot to provide a pleasant and enjoyable interaction experience while also being persuasive. Previous research suggested that the voice is a key element in human communication because it conveys information as well as social cues [11]. Manipulating the voice characteristics, such as pitch, intonation, and gender, can change the people's impression of the robot and the corresponding interaction behaviors [10, 3, 6].

Among these characteristics, the pitch has been found to considerably influence people's communication with a robot. For example, [11] found that a high-pitched robot was perceived as significantly more attractive than a low-pitched one, and its use led to a better overall enjoyment and inter- action quality. In addition, [10] observed that a high-pitched robot was better rated in terms of its appeal as well as the overall interaction quality and enjoyment. Besides the pitch, the user gender and robot voice gender also affect the people's perception of a robot, although the conclusions are inconsistent across studies [8]. Some studies [15] reported that people perceived a robot of the opposite gender more positively, whereas another research [5] concluded that the users evaluated a same-gender robot more positively. Nonetheless, it is assumed that the perceptions of male and female users toward a robot and its voice [4] are different. In terms of the robot gender, [13] demonstrate that people use their knowledge regarding the gender roles when inter- acting with a gendered robot.

Teleoperation is considered to be a promising approach for communication robots because it is able to offer natural human speech communication that most current autonomous robots are not capable of. Performing work using a teleoperated robot allows an increase in the working population, thereby realizing a considerable social contribution [1]. However, most findings pertaining to the effect of voice are those of studies based on autonomous robots. When interacting with teleoperated robots, people may have a different perception of the robots as they may be aware that they are talking to another person behind the robot. For instance, [18] found that the participants had a positive attitude toward a remotely controlled robot. Therefore, it is of importance to evaluate the effect of voice manipulation for teleoperated robots.

To this end, in this study, we investigated whether the voice pitch (original vs. high) and voice gender (male vs. female) affect people's attitudes toward a teleoperated robot and the corresponding decision-making. In addition, we analyzed the influence of the user gender (male vs. female). A teleoperated robot prototype was developed,

and a user experiment was performed. The findings are expected to be beneficial for improving the task performance in teleoperated robot applications.

Method

Task

During the study, 36 Japanese participants (18 males, 18 females) completed one main task, namely, the desert survival task. We used a modified version of the original task proposed by [16], which ensured that the rationale for the suggestion made by the robot was identical regardless of the suggestion (agreement or disagreement) offered.

The participants first filled out a form to indicate their initial item selections after being provided with the task instructions. In particular, the participants needed to make selections from five pairs of items: canvas or tarp, chocolate or water, mirror or compass, flashlight or matches, and knife or pistol. Then, the robot joined the task and began to interact with the participants by facing the participants and raising its right hand. The participants were told that the robot was remotely controlled by another participant who was actually an experimenter. For each pair of items, the robot asked the participants which item they selected and then offered a suggestion regarding their choice by either agreeing or disagreeing with the choice. In particular, the robot always disagreed with the participants' second, fourth, and fifth initial item selections. The participants could then decide whether they wished to change their original choice. This process was repeated for all the five pairs of items. Figure 1 demonstrates the setting of the experiment.

Voice Preparation

In total, we prepared four types of voice recordings for the experiment: (1) normal male voice, (2) normal female voice, (3) high-pitched male voice, and (4) high-pitched female voice.

Teleoperated Robot System

A small humanoid robot called "SOTA" (Vstone Co., Ltd) was used, as shown in Figure 1. The details regarding the design of the system are provided in a previous paper [1]. The robot can rotate in 180◦ and perform motions such as nodding and raising the hand. Because the system is capable of face recognition, the robot could keep facing the participants.

Figure 1: Experiment setting.

Measures

We evaluated the participants' behavior change as well as their attitudes toward the robot. The behavioral measure was basically the change between the initial and final selections of the survival items for which the robot expressed disagreement (second, fourth, and fifth item pairs). However, we explored the data and observed that the distribution of the change in items for the second and fifth item pairs (chocolate vs. water and knife vs. pistol) significantly differed from that for the fourth item pairs (flashlight vs. matches). Specifically, most participants (35 out of 36) made the same initial selections for the second (water) and fifth (knife) item pairs, although their opinions differed in the choice of the fourth item. We presume that there may be a culture bias in the selection of survival items. Thus, we divided the behavioral measure into two variables, disgreement25 and disagreement4, and performed separate statistical analyses.

The attitudinal measures pertained to two structured questionnaires: Godspeed [2] (5-point semantic differential scales) and Trust (modified questionnaires of reliance intentions [14] and trustworthiness [9]; 7-point Likert scales).

One single-item 7-point Likert scale question, "I felt that the robot was teleoperated by someone," was used to explore the relationship between to what extent the participants felt that the robot was teleoperated by a

human operator and the participants' decision-making.

Results

Behavioral Measure

Overall, we found a significant main effect for the user gender [$F(1, 28) = 4.59$, $p < .05$]. The post hoc t-test indicated that females changed their initial selections more often than males did ($p < .05$).

We found a trend of a significant main effect for the user gender [$F(1, 28) = 3.69$, $p = .06$]. The post hoc t-test indicated that females changed their initial selections more often than males did ($p < .05$). We also found a trend of significant interaction effect [$F(1, 28) = 3.69$, $p = .06$] between the voice pitch and user gender. The post hoc comparisons using Tukey's HSD test revealed that females changed their selection more than males, particularly when the robot voice was high-pitched ($p < .05$).

No significant results were observed for the fourth item pairs.

Attitudinal Measure

A significant main effect was observed for the user gender [$F(1, 28) = 4.54$, $p < .05$]. The post hoc t-test indicated that females perceived the robot to be more intelligent compared to that perceived by the males ($p < .05$). A significant interaction effect between the voice pitch and user gender [$F(1, 28) = 4.54$, $p < .05$] was also observed. The post hoc comparisons using Tukey's HSD test suggested that females rated the robot as more intelligent than that considered by the males, particularly when the robot voice was high-pitched ($p < .05$).

Awareness of Teleoperation

A linear regression was performed to predict the change in the survival items based on the participants' awareness of teleoperation. For disagreement25, a significant regression relation was noted [$F(1, 34) = 5.91$, $p < .05$; $R2 = 0.15$], indicating that the participants changed their initial selections more if they had a stronger feeling that the robot was teleoperated. No significant result was found for disagreement4.

Discussion

The current results suggest that the user gender influences the people's attitude toward a teleoperated robot and their decision-making. It appears that females perceive a robot more positively and are easier to persuade compared to males. Overall, the female participants perceived the robot to be more intelligent and changed the initial selections of survival items more often than males did when the robot disagreed with their choice (second, fourth, and fifth item pairs).

Although we did not observe any main effects pertaining to the robot voice gender and pitch on the attitudinal responses toward the robot and change in the items, interaction effects between the voice pitch and user gender were observed. The influence of the user gender was particularly strong when a high-pitched voice was used. Specifically, the perceived intelligence of the robot and change in the items increased in the case of female participants but decreased in the case of male participants when a high-pitched voice was used.

Based on the results, it can be concluded that the use of a high-pitched voice of a teleoperated robot may differently influence the male and female users. In particular, a high-pitched voice may negatively influence males such as in terms of a decrease in the robot's persuasiveness and dependability. Previous findings [11, 10] demonstrated that a robot could be considered more attractive and achieve a higher interaction quality if it spoke in a higher pitch. However, in the case of teleoperated robots, the pitch of the robot's voice must be carefully selected.

The results pertaining to the participants' awareness of teleoperation suggest a plausible explanation for the interaction effects between the voice pitch and user gender. In particular, it was observed that the participants changed their initial selections more often if they had a stronger feeling that the robot was teleoperated. Moreover, it was observed that the average rating for the awareness of teleoperation for female participants increased (from 4.78 to 5) when a high-pitched voice was used whereas the corresponding rating for male participants decreased (from 4.13 to 3.73).

Such a finding is consistent with the effects of high-pitched voice on males and females. Thus, we speculate that when a high-pitched voice was used, the female

participants tended to perceive the robot as being teleoperated whereas the male participants treated the robot as being autonomous. The researchers in [18] indicated that people may prefer a remotely controlled robot as they intuitively find the robot to be more secure and reliable when a human is the operator. Similarly, our participants likely considered the suggestions pertaining to the survival items to be more reliable if they felt that these suggestions were provided by a human operator compared to those from an autonomous robot. As a result, the female participants accepted the robot's suggestions and changed their choice of items when a high- pitched voice was used; however, the opposite results were observed for males.

In our study, we divided the behavioral measure into two variables, disgreement25 and disagreement4, because most participants made the same initial selections for the second (water) and fifth (knife) item pairs but not for the fourth one (flashlight vs. matches). We presumed that the participants, owing to the culture and education system in Japan, hold strong preferences toward the selection of the second and fifth item pairs but no obvious preference toward the fourth one. Interestingly, significant results were found for disagreement25 but not in the case of disagreement4. In other words, the effects of high-pitched voice only appeared when the robot disagreed with the initial selections of items which the participants had strong preferences for.

We speculate that, for the fourth item pairs, which the participants did not exhibit an obvious preference for, the participants considered the content of the information (i.e. the benefit of each item) more important than the judgment made by the robot. However, for the second and fifth item pairs, the results regarding the change in the items corresponded to the persuasiveness of the robot. This could be considered as an important implication for future teleoperated robot applications, for example, for providing product recommendations. For the products that people are not familiar with, the robots may focus more on providing information about the products; however, for the products that people are familiar with, robots may need more expressivity.

One limitation of this work is that the number of participants was small, which led to limited findings. Another limitation is that we did not investigate the effects of low-pitched voice. Therefore, in future research, the number of participants and types of voices may be increased.

Reference:

[1] Jun Baba, Sichao Song, Junya Nakanishi, Itaru Kuramoto, Kohei Ogawa, Yuichiro Yoshikawa, and Hiroshi Ishiguro. 2020. Teleoperated Robot Acting Autonomous for Better Customer Satisfaction. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. ACM, (in press). DOI: http://dx.doi.org/10.1145/3334480.3375212

[2] Christoph Bartneck, Elizabeth Croft, and Dana Kulic. 2008. Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots. (2008).

[3] Rebecca Cherng-Shiow Chang, Hsi-Peng Lu, and Peishan Yang. 2018. Stereotypes or golden rules? Exploring likable voice traits of social robots as active aging companions for tech-savvy baby boomers in Taiwan. Computers in Human Behavior 84 (2018), 194–210. DOI: http://dx.doi.org/10.1016/j.chb.2018.02.025

[4] Charles R Crowelly, Michael Villanoy, Matthias Scheutzz, and Paul Schermerhornz. 2009. Gendered voice and robot entities: perceptions and reactions of male and female subjects. In 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 3735–3741. DOI: http://dx.doi.org/10.1109/IROS.2009.5354204

[5] Friederike Eyssel, Laura De Ruiter, Dieta Kuchenbrandt, Simon Bobinger, and Frank Hegel. 2012. 'If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In 2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, 125–126. DOI: http://dx.doi.org/10.1145/2157689.2157717

[6] Jaeyeol Jeong and Dong-Hee Shin. 2015. It's not What It Speaks, but It's How It Speaks: A Study into Smartphone Voice-User Interfaces (VUI). In International Conference on Human-Computer Interaction. Springer, 284–291. DOI: http://dx.doi.org/10.1007/978-3-319-20916-6_27

[7] Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. 2010. A communication robot in a shopping mall. IEEE Transactions on Robotics 26, 5 (2010), 897–913. DOI: http://dx.doi.org/10.1109/TRO.2010.2062550

[8] Dieta Kuchenbrandt, Markus Häring, Jessica Eichberg, Friederike Eyssel, and Elisabeth André. 2014. Keep an eye on the task! How gender typicality of tasks influence

human–robot interactions. International Journal of Social Robotics 6, 3 (2014), 417–427. DOI: http://dx.doi.org/10.1007/s12369-014-0244-0

[9] Roger C Mayer and James H Davis. 1999. The effect of the performance appraisal system on trust for management: A field quasi-experiment. Journal of applied psychology 84, 1 (1999), 123.

[10] Andreea Niculescu, Betsy van Dijk, Anton Nijholt, Haizhou Li, and Swee Lan See. 2013. Making social robots more attractive: the effects of voice pitch, humor and empathy. International journal of social robotics 5, 2 (2013), 171–191. DOI: http://dx.doi.org/10.1007/s12369-012-0171-x

[11] Andreea Niculescu, Betsy Van Dijk, Anton Nijholt, and Swee Lan See. 2011. The influence of voice pitch on the evaluation of a social robot receptionist. In 2011 International Conference on User Science and Engineering (i-USEr). IEEE, 18–23. DOI: http://dx.doi.org/10.1109/iUSEr.2011.6150529

[12] Roberto Pinillos, Samuel Marcos, Raul Feliz, Eduardo Zalama, and Jaime Gómez-García-Bermejo. 2016. Long-term assessment of a service robot in a hotel environment. Robotics and Autonomous Systems 79 (2016), 40–57. DOI: http://dx.doi.org/10.1016/j.robot.2016.01.014

[13] Aaron Powers, Adam DI Kramer, Shirlene Lim, Jean Kuo, Sau-lai Lee, and Sara Kiesler. 2005. Eliciting information from people with a gendered humanoid robot. In ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005. IEEE, 158–163. DOI: http://dx.doi.org/10.1109/ROMAN.2005.1513773

[14] F David Schoorman, Roger C Mayer, and James H Davis. 2007. An integrative model of organizational trust: Past, present, and future. (2007). DOI: http://dx.doi.org/10.5465/amr.2007.24348410

[15] Mikey Siegel, Cynthia Breazeal, and Michael I Norton. 2009. Persuasive robotics: The influence of robot gender on human behavior. In 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2563–2568. DOI: http://dx.doi.org/10.1109/IROS.2009.5354116

[16] Leila Takayama, Victoria Groom, and Clifford Nass. 2009. I'm sorry, Dave: i'm afraid i won't do that: social aspects of human-agent conflict. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2099–2108. DOI: http://dx.doi.org/10.1145/1518701.1519021

[17] Meg Tonkin, Jonathan Vitale, Sarita Herse, Mary-Anne Williams, William Judge, and Xun Wang. 2018. Design methodology for the ux of hri: A field study of a commercial social robot at an airport. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. ACM, 407–415. DOI: http://dx.doi.org/10.1145/3171221.3171270

[18] Astrid Weiss, Daniela Wurhofer, Michael Lankes, and Manfred Tscheligi. 2009. Autonomous vs. tele-operated: How people perceive human-robot collaboration with HRP-2. In 2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, 257–258. DOI: http://dx.doi.org/10.1145/1514095.1514164