Speech Related Accessibility Issues in Social Robots

Anouk van Maris
anouk.vanmaris@uwe.ac.uk
Bristol Robotics Laboratory
Bristol, United Kingdom

Sanja Dogramadzi
sanja.dogramadzi@uwe.ac.uk
Bristol Robotics Laboratory
Bristol, United Kingdom

Nancy Zook
nancy.zook@uwe.ac.uk
Department of Health and Social Sciences, UWE
Bristol, United Kingdom

Matthew Studley
matthew.studley2@uwe.ac.uk
Bristol Robotics Laboratory
Bristol, United Kingdom

Alan Winfield
alan.winfield@brl.ac.uk
Bristol Robotics Laboratory
Bristol, United Kingdom

Praminda Caleb-Solly
praminda.caleb-solly@uwe.ac.uk
Bristol Robotics Laboratory
Bristol, United Kingdom

ABSTRACT

This work describes an incidental finding from a longitudinal Human-Robot Interaction study that was investigating whether a robot showing emotions during interactions with older adults was perceived differently than to a robot that did not display emotions during the interaction. During this study we noted that some older adults found it hard to understand what the robot was saying, regardless of the volume of speech generated by the robot. The fact that they did not have problems in understanding the researcher led us to investigating this accessibility-related issue in more depth. This paper describes the implications of this finding and recommendations on how to approach future work.

KEYWORDS

accessibility; social robot; older adults; human-robot interaction

ACM Reference Format:
Anouk van Maris, Sanja Dogramadzi, Nancy Zook, Matthew Studley, Alan Winfield, and Praminda Caleb-Solly. 2020. Speech Related Accessibility Issues in Social Robots. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20 Companion), March 23–26, 2020, Cambridge, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3371382.3378379

1 INTRODUCTION

While a lot of research and best practice on the design of accessible user interfaces exists in the field of Human-Computer Interaction, with even legislature regarding web accessibility [4], there is not as much accessibility related research in Human-Robot Interaction as yet. Our research relates to investigating the efficacy of assistive robots to support people with ageing-related impairments to live independently for as long as possible. In addition to exploring how robots can provide support for activities of daily living, our research is also concerned with the multitude of ethical aspects regarding the use of assistive robots. One of these issues is emotional deception, the impact of people perceiving a social robot as being sentient because it expresses emotions and behaviour that could be interpreted as it being capable of having emotions. As part of this research, we have been conducting longitudinal research studies with older adults. Ageing results in increased frailty, including sensory, cognitive and physical impairments resulting in a range of accessibility issues. In addition, the severity of the ageing-related impairments changes over time so a system that was once effective, can later become unsuitable due to progressive decline of the person's condition [2].

Ensuring effective interaction therefore requires a clear understanding of the impact of ageing-related impairments, and how these might be ameliorated by improving the design of the system by considering multi-modal interaction utilising a range of communication channels [6]. Improved understanding of the impact of these on the quality of the interaction and on the person's experience can result in enhancing the effectiveness of the assistive solutions. In our research, where we used the Pepper platform from SoftBank Robotics [5], we set out to discover the impact of interaction of different robot behaviours in situ, but in examining our results noted that some people were experiencing difficulties in understanding the speech generated by the robot. Our research was conducted in supported retirement village settings over a period of several weeks. The problems with accessibility were not something that we had noted under laboratory conditions in other studies with the same Pepper platform where participants (also older adults) were invited to the lab.

In this paper we provide our findings related to accessibility, discuss our approach for addressing this in a subsequent study and consider other issues which researchers using similar social robot platforms might want to consider.

2 METHODOLOGY

The study that provided the incidental findings continues on earlier work involving the effects of emotional deception by a social robot on its users, where emotional deception is defined as misrepresenting one's emotional state [3]. This previous work (in which Pepper was used as well) showed that there was not a strong effect of a robot's (non-) emotional behaviour on their perception of the interaction, acceptance of the robot, or their mood [8],[1]. The follow-up study described in this paper took into account time, as the study was longitudinal and participants had eight interactions with the robot over a period of six weeks. The nature of the interactions was didactic as the robot informed the participants on the Seven

HRI '20 Companion, March 23–26, 2020, Cambridge, United Kingdom

Wonders of the World. The robot would give them information about a specific wonder each session, ask them what they knew about that wonder and whether they had visited it. Responses of the robot were pre-programmed for consistency, but the robot was manually prompted to continue the interaction after asking the participant a question by the experimenter. It was decided to wizard this part of the interaction as speech recognition is not always reliable and may therefore result in a frustrating experience for the participants if they are required to repeat themselves. The study was conducted at two retirement villages for older adults. There is a growing body of research looking into the use of robots to support an ageing population, where there are not enough trained care staff to offer a high level of support. As older adults, particularly those with ageing-related cognitive impairment and frailty can be quite vulnerable in terms of their situation and level of need, there is concern from an ethical perspective that this demographic may be more sensitive to displays of emotion by a social robot. As this study was performed at the retirement villages and not in a lab setting, this both provided access to a more diverse participant group and not only people who are willing and fit enough to travel to the location where the experiment is being conducted. Additionally, in order to ensure ecological validity of the experiments, we were also mindful that participants may feel more comfortable in their home environment and therefore respond differently to the robot.

3 RESULTS

In total 14 participants (9 male, 5 female) from two residential villages took part in this longitudinal experiment. At the first retirement village, four participants (3 male, 1 female) completed the experiment. These participants interacted with the robot without it providing subtitles of what it was saying on its tablet. It was found that participants found it difficult to understand the robot. They reported the volume of the robot was good, which indicates that the pronunciation of the robot may have caused the issue. As this was found early on in the experiment, subtitles were added for the second retirement village where ten participants interacted with the robot (6 male, 4 female). Nearly all participants of this retirement village claimed to fully understand the robot. However, it was observed by the experimenter that sometimes participants would respond to the robot when they finished reading the subtitles on the tablet, even though the robot had not finished speaking yet. The interactions between the participants and the robot were video-recorded. Initial transcriptions of these recordings show that participants ask the robot more often to repeat itself when no subtitles are provided, as it was only asked once when subtitles were provided as compared to 15 times when subtitles were not provided.

All four participants in the first setting asked the robot to repeat itself, indicating that it was not due to one participant being hard of hearing. This strengthens the observation that participants had trouble understanding what the robot was saying. The experimenter also noticed that participants would always blame themselves for not understanding the robot, claiming that they either needed a hearing aid or that the hearing aid they were wearing must be the problem. This again shows the importance of a social robot being accessible to all its users, as interactions to benefit the user should never result in the user feeling bad about themselves.

We found that not providing subtitles results in relatively richer interactions, with participants commenting more often on what the robot was saying without it asking a question when subtitles were not provided (M comments per interaction = 20, SD = 5.57 over all 8 interactions) in comparison to when subtitles were provided (M = 12.13, SD = 12.01 over all 8 interactions). However, the number of words that participants use during the interaction appear to be fewer for participants that had no subtitles (M = 40, SD = 15.88 per session) as compared to participants that did have subtitles (M = 45, SD = 10.85 per session). Perhaps this is due to the fact that participants with subtitles have a reminder of the question or comment from the robot and therefore are able to provide more elaborate answers.

4 DISCUSSION

Even though the goal of this experiment was to investigate the effect of displayed emotions by a robot, there was an incidental finding relating to accessibility needs that drew our attention. While providing subtitles resulted in an improvement, it must also be noted that older adults might also have visual impairments, or the robot screen might not always in their field of view or at a suitable viewing angle. Additionally, in a home environment, variations in lighting conditions could also lead to people having problems with reading. It is therefore necessary to consider how part of the system intelligence should also take into account being able to detect automatically when people are having difficulties with understanding or are making errors in their response. The content and amount of interaction could also be used as key indicators of issues as seen from our results. Even though the providing subtitles appeared to reduce the issue of participants not understanding the robot, it must be noted that it was an assumption that the pronunciation was the issue and other factors may have been of influence as well. This will be investigated further in future work. Additionally, future work could consider use of gesture or colour changes to enhance or emphasise the content, which could improve comprehension and understanding [7].

ACKNOWLEDGMENTS

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721619 for the SOCRATES project. We would also like to thank all participants for taking part in this experiment.

REFERENCES

[1] Laura Bishop, Anouk van Maris, Sanja Dogramadzi, and Nancy Zook. 2019. Social robots: the influence of human and robot characteristics on acceptance. *Paladyn Journal of Behavioral Robotics* 10, 1 (2019), 346–358.
[2] Praminda Caleb-Solly. 2016. Person-environment interaction. (2016).
[3] Ingrid Smithey Fulmer, Bruce Barry, and D Adam Long. 2009. Lying and smiling: Informational and emotional deception in negotiation. *Journal of Business Ethics* 88, 4 (2009), 691–709.
[4] Jonathan Lazar. 2019. Web Accessibility Policy and Law. In *Web Accessibility*. Springer, 247–261.
[5] Amit Kumar Pandey and Rodolphe Gelin. 2018. A mass-produced sociable humanoid robot: pepper: the first machine of its kind. *IEEE Robotics & Automation Magazine* 25, 3 (2018), 40–48.
[6] Tanya Stivers and Jack Sidnell. 2005. Introduction: multimodal interaction. *Semiotica* 2005, 156 (2005), 1–20.

[7] Ayano Sueyoshi and Debra M Hardison. 2005. The role of gestures and facial cues in second language listening comprehension. *Language Learning* 55, 4 (2005), 661–699.

[8] Anouk van Maris, Nancy Zook, Praminda Caleb-Solly, Matthew Studley, Alan Winfield, and Sanja Dogramadzi. 2018. Ethical considerations of (contextually) affective robot behaviour. In *Hybrid Worlds: Societal and Ethical Challenges - Proceedings of the International Conference on Robot Ethics and Standards (ICRES 2018)*. CLAWAR Association Ltd, 13–19.