Robot Response Behaviors To Accommodate Hearing Problems*

Jered Vroon[a], Jaebok Kim[a], and Raphaël Koster[b]

[a]*University of Twente, Human Media Interaction*
[b]*MADoPA*

Abstract. One requirement that arises for a social (semi-autonomous telepresence) robot aimed at conversations with the elderly, is to accommodate hearing problems. In this paper we compare two approaches to this requirement; (1) moving closer, mimicking the leaning behavior commonly observed in elderly with hearing problems, (2) turning up the volume, which is a more mechanical solution. Our findings with elderly participants show that they preferred the turning up of the volume, since they rated it significantly higher.

Keywords: Telepresence robot, Hearing problems

INTRODUCTION

What behavior is appropriate for a social robot will depend on the context in which it is to function. For example, for a robot that helps lifting people out of bed it is necessary to get intimately close, while for a telepresence robot such intimate distances probably are less appropriate. An important aspect of this context are the specific individual needs of the users.

Elderly with hearing problems are one such user group that places its own requirements on the behavior of social robots. Hearing problems have a high prevalence among elderly [1,2]. Taking hearing problems into account could thus be a good contribution to any robot that is to communicate through audio with elderly, such as for example (semi-autonomous) telepresence robots.

One way to handle hearing problems is by mimicking the 'leaning' behavior commonly observed in this user group, where people actively lean in to intimate distances during conversations [3,4]. Similarly, a social conversation robot could also reciprocate such leaning behavior by moving closer.

An alternative would be to instead change the volume settings of the robot. Though in a way less human-like, this could be equally (or more) effective in resolving the hearing problems.

The aim of the reported experiment is to investigate with elderly participants which of these two response behaviors they might prefer a semi-autonomous telepresence robot to show.

METHOD

To investigate the effect of the different response behaviors, we set up a within subject experiment [*no response* x *move closer* x *turn up volume*] as part of a larger evaluation session for the Teresa project*. In each session one participant (the **Visitor**) sat in a remote location and used the robot in another room to interact with one or two other participants in the same room as the robot (the **Interaction Target(s)**). We used a Giraff[1] telepresence robot. A possible limitation is that the speaker is located in its base, not its 'head'.

Procedure

The Interaction Target(s) were seated behind a table, with the robot on the other end of it at a distance of approximately 1.5m. To ensure that hearing problems would arise, the volume of the robot had been turned down to a barely audible level. An experimenter explaining the procedure sat with the Interaction Target(s) during the experiment.

To make the conditions more comparable, the experiment started with a full briefing on the aim and the procedure of the experiment. After this, there were three trials in which participants had a brief conversation with each other that was terminated after about two minutes by the experimenter. In each of these trials, as soon as the Interaction Target(s) expressed having hearing problems or after approximately one minute, a Wizard of Oz showed one of the three response behaviors in counterbalanced order. For *'no response'*, no behavior was shown. For *'move closer'*, the robot approached the Interaction Target(s) to a distance of around 0.8m. For *'turn up volume'*, the volume settings were turned up a bit, which was also visible in the interface. To ensure functional comparability, none of these changes was sufficient to completely resolve all hearing problems. At the end of each trial, the robot was returned to its initial position and volume setting. The experiment was concluded with a brief (paper) questionnaire.

Task

To stay close to the intended use of the robot, the task of our participants was to have a conversation. For this, we asked them to discuss questions of the Proust questionnaire[2]. Specifically, we asked the Interaction Target(s) to read out self-selected questions and the Visitors to discuss what they thought the Interaction Target would answer.

Measurements

At the end of the interactions, all participants were given a brief questionnaire. Three items asked them to indicate their most and least favorite response behavior and to rate all response behaviors on a scale of 1-10.

*This work has been supported by the European Commission under contract number FP7-ICT-611153 (TERESA), http://teresaproject.eu

[1] http://www.giraff.org/?lang=en

[2] http://fr.wikipedia.org/wiki/Questionnaire_de_Proust

Table 1. Descriptive statistics for the ratings given to the three different response behaviors.

Response behavior	N	Mean	Percentiles				
			Min	**Q25**	**Q50**	**Q75**	**Max**
No response	18	3.000	0.0	1.5	3.0	4.25	9.0
Move closer	17	6.167	0.0	5.0	6.5	8.0	9.0
Turn up volume	18	8.235	6.0	7.5	8.0	9.5	10.0

Table 2. Number of times the different qualities were checked as being most influential in giving the ratings (total = 54).

Agréable (Pleasant)	**Sensible** (Sensitive)	**Amical** (Friendly)	**Utile** (Helpful)	**Affable** (Affable)	**Sympathique**(Likeable)	**Accessible**(Approachable)	**Sociable** (Sociable)	**Intelligent** (Intelligent)	**Organisé** (Organized)	**Expert** (Expert)	**Réfléchi** (Thorough)	**Efficace** (Effective)	**Approprié**(Appropriate)	**Attentif** (Attentive)
2	0	2	9	2	1	4	4	10	2	0	4	5	4	5

One item asked them to indicate which three qualities of the robot were most influential in their ratings, based on items for warmth and competence [5] (see Table 2 for the qualities). The last 5 items considered demographics (age, gender, hearing problems, use of hearing aids, relationship with the other participant(s)).

We recorded the interactions on video and using robot-mounted sensors. The interface as seen by the Visitor was recorded using screen capture software.

Participants

We had 18 French speaking participants (13 female, 4 male, 1 undisclosed), in six pairs and two trios, all with a prior relation (e.g. friends, family). Participant were aged between 60 and 91 (mean age 74). Hearing loss was reported by 7 participants. In one pair, a 10-year old grand-child also joined as Interaction Target, but was excluded from analysis.

FINDINGS

Summaries of our main findings can be found in Tables 1 and 2. Twelve participants preferred the *'turn up volume'* behavior, the other six preferred *'move closer'* instead. The ratings of these behaviors matched those preferences for 89% of the participants, though many asked for clarification of the rating questions.

Since the rating of the response behaviors was not normally distributed (Shapiro-Wilk, p=0.135, p=0.039*, p=0.053) we ran a Friedman test, which found a significant difference in rating ($\chi^2(2)=25.344$, p=0.000*). We did a post hoc analysis with a Wilcoxon signed-rank test (significance level 0.017, with Bonferroni correction). The ratings for *'move closer'* were significantly higher than those for *'no response'* (Z=-2.917, p=0.004*). The ratings for *'turn up volume'* were significantly higher than both those for *'no response'* (Z=-3.628, p=0.000*) and those for *'move closer'* (Z=-2.462, p=0.014*).

This analysis made the simplifying assumption that the participants can be treated as independent comparable measurements, despite being in the same group and having one of two roles (Visitor/Interaction Target). A series of Pearson's Chi-square test found no significant correlations of either group or role with the ratings, which supports this assumption. The aforementioned significant differences all hold when looking at the Interaction Targets only (N=10), only the difference in rating for *'turn up volume'* and *'move closer'* is no longer significant (Z=-1.364, p=0.172).

CONCLUSIONS AND DISCUSSION

We have compared three ways in which a semi-autonomous telepresence robot could respond to hearing problems. We found high ratings for *'turn up volume'*, significantly surpassing the ratings for *'move closer'*. Both of these were rated significantly higher than *'no response'*. There do seem to be further individual differences, as one third of the participants instead preferred the *'move closer'* behavior. We only used general ratings for this, but our participants most commonly indicated to have based their judgement mostly on the qualities 'Intelligent' and 'Helpful'. Note that these findings need not translate to other settings, e.g. *'turn up volume'* may be perceived as less appropriate if the noise could disturb others.

Overall, our findings demonstrate that trying to accommodating hearing problems is a desirable feature in this setting. A general approach like turning up the volume when required could work in general cases. If possible, a more personalized solution could also/instead move closer if the user would so prefer.

ACKNOWLEDGEMENTS

The authors wish to thank Khiet Truong, Gwenn Englebienne and Vanessa Evers for their input during the design of the experiment, and Alexandre Duclos for being the Wizard of Oz.

REFERENCES

1. N. Blevins, D. Deschler and L. Park. "Presbycusis". Retrieved September 15, 2014, from *www.uptodate.com/contents/presbycusis*
2. K.J. Cruickshanks, et al. "Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin the epidemiology of hearing loss study." *American Journal of Epidemiology* **148**(9), 879-886 (1998).
3. J. Vroon, G. Englebienne and V. Evers. "D3.1 Normative Behavior Report". Teresa project deliverable, 2014. Available from *teresaproject.eu/project/deliverables/*
4. J.D. Webb and M.J. Weber. "Influence of sensory abilities on the interpersonal distance of the elderly". *Environment and behavior* **35**(5), 695-711 (2003).
5. K. Bergmann, F. Eyssel and S. Kopp. "A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time." *Intelligent Virtual Agents*. Springer Berlin Heidelberg, 2012.