

Using robots to study speech – interactional processes and speech functions

Kerstin Fischer

Department of Design and Communication, University of Southern Denmark,
Denmark

<https://doi.org/10.36505/ExLing-2017/08/0001/000303>

Abstract

In this paper, I present some scenarios in which human-robot interaction can be carried out to study linguistic features and their interactional functions, as well as complex interactional processes.

Key words: speech, language, human-robot interaction, methodology

Introduction

Robots are a great methodological resource for linguistic and interactional research since they can be completely controlled (unlike human confederates) and thus exhibit identical behaviors across participants and conditions. Furthermore, people do not hold particular expectations towards robots, unlike towards other people, which allows us to experiment with a broad range of behaviors.

In the most trivial case, we can use robots in a variant of the matched guise technique (Gardner & Lambert 1972) to study people's attitudes towards certain linguistic features and their interpersonal functions; for instance, we can use them to study the functions of rising and falling intonation contours, speaking styles, durational features, the use of hesitation markers, types of feedback signals and so on. While the traditional matched guise experiment requires that there are several filler items between the two target stimuli that make people forget what they have just heard, in human-robot interaction, we can match the two sound files with different robots, or even easier, with two videos of different robots in a questionnaire. Besides this very simple and obvious use of robots in linguistic research, robots can also be a resource to study complex interactional mechanisms at the intersections between cognitive and social processes, which I will show below.

The next section presents some human-robot interaction scenarios in which linguistic research can be carried out. Then, I focus on one major interactional feature: recipient design, i.e. the processes involved in how we adjust our speech to our particular addressees, for instance, to a child, a foreigner or a robot. I illustrate how we can investigate complex socio-cognitive phenomena by means of human-robot interaction experiments.

Scenarios for linguistic human-robot interaction research

In the following, I describe some human-robot interaction scenarios that have already been successfully employed for linguistic research. While obviously many different scenarios are possible, the following are the most common ones.

Video-based studies

Experiments with robot videos are easy to set up and thus especially attractive for students; for instance, Aarstrup (2014) investigated how people respond to a greeting if it is delivered with different intonation contours. Previous linguistic work had suggested that, for instance, "a fall indicates the speaker's dominance in knowing and telling something, in telling someone what to do, and in expressing their own feelings; in contrast, a rise indicates a speaker's deference to the addressee's knowledge, their right to decide, and their feelings" (Tench 1996: 105). Wells (2006) furthermore holds the greeting '*hi*' not to be used with rising intonation.

Aarstrup used videos of three different robots, in which the robots make a waving gesture or produce a small bow, and synthesized *hi* and *hello* with a free text-to-speech system. She then manipulated the sound files using praat (Boersma 2001) to create stimuli with different intonation contours. Altogether, the experiment uses two lexical items (*hello* vs. *hi*), three robots (Nao, Care-O-bot and Robosapien) and four different intonation contours (fall, rise, fall-rise and flat). The videos were then matched with the different sound files and distributed over four questionnaires so that each participant only got to see each robot once (n=120). Her results reveal that non-native speakers do not respond to the differences in intonation contours or lexical items at all. In contrast, native speakers of English rate the robots significantly different concerning friendliness, assertiveness and engagement depending on the intonation contours. However, these effects differ for the different lexical items, and the apparently non-conventional *hi* with rising intonation contour was in fact rated as most engaging.

Thus, matching videos with different speech stimuli provides a simple tool to investigate interpersonal effects of particular linguistic features.

One-on-one human-robot interactions

One-on-one interaction is the most common form of human-robot interaction experiments. The robot's speech in this scenario can be pre-recorded, pre-synthesized or synthesized online using a text-to-speech system. The range of robots used in this scenario is principally without limits, but if the robot exhibits no human-like characteristics at all, speech may not be the most

plausible mode of communication. All speech phenomena that can be tested in video-based scenarios can also be studied in this scenario, plus most interactional features, such as the timing of utterances and the appropriateness of utterances in an interactional discourse context; in addition, properties of speech can be studied in relation to other modalities, such as gaze or robot movement, and their timing with respect to each other (e.g. Jensen et al., in press).

One human – two similar robots

In order to test behavioral effects of speech characteristics, Andrist et al. (2013, 2015) developed a scenario in which the participant is interacting with two similar robots. The two robots take turns presenting information to the participant; for instance, they both make suggestions for places to visit, from which the participant has to choose one. Crucially, the robots' speech differs with respect to the speech characteristic(s) under consideration. Andrist et al. (2013) investigate the effects of certain rhetorical strategies indicating expertise, Andrist et al. (2015) use the same scenario to study the interactions between clues to expertise and linguistic variety, in their case Standard Arabic versus the local Lebanese dialect. Their behavioral measure is how often participants follow the recommendations of the robot that used the strategy under consideration.

Using robots to investigate interaction processes

Recipient design is a central feature of all speech, but it becomes most obvious when people are talking to interaction partners who are 'peculiar' in some way or other, for instance, children, foreigners or dogs. While it is thus a common place that we adjust to our communication partners (cf. Brennan & Schober 2003), it is less clear how we do it.

In Fischer (2016), I carried out several human-robot interaction experiments to identify the contributions of a) a partner model, i.e. some kind of cognitive representation of the partner, of b) automatic adjustments to the partner's speech based on alignment (Pickering & Garrod 2004), and of c) the effects of the partner's feedback signals for the choices speakers make. For example, one set of experiments investigated how the way people talked to a robot changed depending on whether the robot produced speech itself. Because the linguistic output of the robot could be controlled completely, I could nail down in detail how much vocabulary and linguistic structures from the robots' utterances people aligned with and what else their linguistic behavior was informed by, disentangling the effects of partner models and automatic alignment. The results show that people's linguistic choices depend more on their partner model than on the re-use of the robot's linguistic behavior; in fact, the amount of 'automatic' alignment (priming)

was found to crucially depend on people's partner models. Similarly, the effects of the communication partner's (i.e. the robot's) contributions could be determined; while the robots' appearance was not found to have much influence, people use all kinds of behavioral clues to revise and update their partner models of the robots, which in turn determines their linguistic behavior.

To sum up, human-robot interactions are highly useful for all kinds of linguistic research, including the analysis of highly complex interactional phenomena.

Acknowledgements

The research reported on here was conducted within the project *Improving Second Language Pedagogy at the Prosody-Pragmatic Interface by Using Human-Robot Interaction*. The project is financed by the Danish Council for Independent Research.

References

- Aarstrup, M.I.C. 2014. Greetings and Intonation Contours in Human-Robot Interaction. BA Thesis, University of Southern Denmark.
- Andrist, S., Spannan, E., and Mutlu, B. (2013). Rhetorical Robots: Making Robots More Effective Speakers Using Linguistic Cues of Expertise. In *Proceedings of HRI '13*, pp. 341-348.
- Andrist, S., Zidee, M., Boukaram, H., Mutlu, B., and Sakr, M. (2015). Effects of Culture on the Credibility of Robot Speech: A Comparison between English and Arabic. In *Proceedings of HRI '15*, pp. 157-164.
- Boersma, P. 2001. Praat: A system for doing phonetics by computer. *Glott International* 4, 341-345.
- Fischer, K. 2016. Designing Speech for a Recipient. Partner Modeling, Alignment and feedback in so-called Simplified Registers. Amsterdam: John Benjamins.
- Gardner, R.G., Lambert, W. 1972. *Attitudes and motivation in a second-language learning*. Rowley Massachusetts: Newbury House Publishers.
- Jensen, L.C., Fischer, K., Bodenhausen, L. Suvei, S.D. in press. The Synchronization of Multimodal Robot Behaviors during Human-Robot Collaboration. *Proceedings of Ro-Man'17*, Lisbon, Portugal.
- Pickering, M.J., Garrod, S. 2004. Towards a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27, 169–225.
- Schober, M.F., Brennan S.F. 2003. Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher, S. R. Goldman (Eds.), *Handbook of Discourse Processes*, pp. 123–164. Hillsdale: Lawrence Erlbaum.
- Tench, P. 1996. *The Intonation Systems of English*. London: Cassell.
- Wells, J. C. 2006. *English Intonation: an introduction*. Cambridge: Cambridge University Press.