

Studying Language Attitudes Using Robots

Kerstin Fischer

Department of Design and Communication
University of Southern Denmark
Sonderborg, Denmark
kerstin@sdu.dk

Oliver Niebuhr

Department of Industrial Electronics
University of Southern Denmark
Sonderborg, Denmark
olni@sdu.dk

ABSTRACT

Research on language attitudes concerns the identification of the beliefs people hold towards speakers of a particular variety (for instance, a dialect) or towards speakers with a foreign accent. While researchers have been very creative in finding methods for determining speaker attitudes towards their own and others' linguistic productions, robots provide an excellent methodological tool to study language attitudes. We illustrate this methodology on the perception of transfer of speech melody from one's mother tongue to a second language. Our results show effects of such transfer on the perception of the respective speaker; for instance, Danish speakers may be perceived as dominant when transferring their intonation contours to German, whereas Germans may appear formal when transferring their speech melody into Danish.

CCS CONCEPTS

• Human-centered computing~User studies • Social and professional topics~Cultural characteristics

KEYWORDS

robots as methodological tools, linguistic varieties, intonation

1. Introduction

Many features of our linguistic productions may have direct interpersonal effects, i.e. they may have an impact on how we are perceived as a person (e.g. Giles et al. 1981). Sociolinguists address such issues in language attitude research, determining a) what linguistic features are relevant for how speakers are perceived, and b) what attitudes are connected to those linguistic features.

The most common method for studying language attitudes is the matched guise technique, originally introduced by Lambert et al. (1960). In this methodology, a bilingual speaker produces the same utterances in the two languages or dialects under consideration; participants are then asked to rate the speaker based on his or her voice concerning all suspected attitudes, such as his or her

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

submitted to HRI '20 Companion, March 23–26, 2020, Cambridge, United Kingdom
© 2020 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-7057-8/20/03.

intelligence, social background, education etc. To disguise the fact that participants hear the same person twice, several other 'distractor' utterances from other speakers are played in between so that participants do not notice that they have heard the same voice before. The same speaker is used in order to ensure that the voice of the speaker and the potential clues it may give to the speaker's personality do not introduce confounding factors but are identical across the target stimuli. While the method is not without problems, it is still commonly used today (e.g. Soukop 2013).

Now, using robots has several advantages: Since robots can serve as embodied interaction partners, they can take over human social roles, so that it is possible to study the effects of linguistic variables in interactional, and even behavioral, contexts (Fischer 2016; Fischer et al. 2019). Because robots can be both manipulated and completely controlled, they can be trusted to deliver the same behavior towards each participant. Furthermore, because two robots of the same kind look identical, efforts to disguise the speaker and distractor items are superfluous.

Andrist et al. (2013, 2015) first used robots to study the effects of certain linguistic features on the perception of the user. In their experiments, they had robots take turns in presenting information about sights to see, from which participants could then choose. The robots' presentations differ in the kinds of clues to expertise that the robots were employing. While Andrist et al. (2013, 2015) switch the robot presenting, for instance, the expert versus non-expert stimuli during the interactions, this is obviously not possible when robots are speaking in different voices or varieties. We thus change the order of presentation only between subjects, so that each participant hears each robot consistently speaking in one way. We then ask people to rate the robots' personalities in comparison. Robots taking turns in presenting linguistic stimuli can thus provide an alternative to studying language attitudes using the matched guise technique or similar methods.

In the following, we illustrate the methodology using the transfer of aspects of speech melody from one's mother tongue into a second language.

2. Example Study: Procedure

The study of intonation transfer involves some representation of the common intonation patterns in the source and in the target

language. However, identifying what is ‘typical’ can be problematic since intonation, i.e. speech melody, depends on many contextual factors, and to some extent also on the speaker him- or herself. To circumvent these problems, we decided for one specific situation both for the elicitation of the stimuli and for the experiment itself, namely asking questions in a questionnaire. In order to elicit naturally occurring speech melodies in the three languages under consideration, namely English, German and Danish, we recorded native speakers of Danish, German and English asking the demographic questions of a questionnaire. Then we manipulated the intonation contours of one file by imposing the intonation contour of the other file using PRAAT (Boersma 2001). We chose question intonation for the study since yes/no-questions in English and German typically end in a rising tone (cf. Bolinger & Bolinger 1989; Niebuhr 2015), whereas questions in Danish end in a level tone (Grønnum 2007). Figure 1 provides an example of the differences in intonation concerning the question ‘how old are you?’. In the manipulation, we imposed the final level tone from the Danish onto the German utterance with the original rising tone.

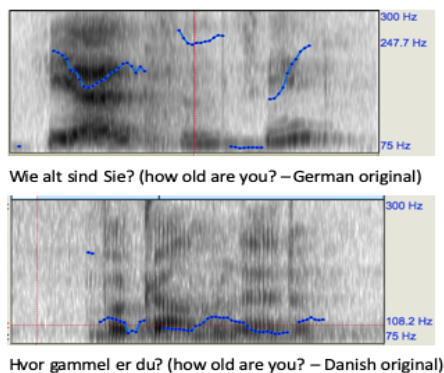


Fig. 1: The intonation contours of the original German and Danish questions ‘how old are you?’ with rising final tone in German and level contour in Danish

The manipulation was carried out in a team of linguists and phoneticians to ensure the naturalness of the resulting questions.

In an online survey, two Keepon robots (Fig. 2) took turns in asking the questions, which the participants had to respond to. One Keepon used the original intonation, the other used the intonation from another language. For example, German participants heard three original German questions and three German questions of the same speaker with the manipulated intonation contours. Afterwards, they had to decide which of the two Keepon robots was more engaging, friendly, polite, dominant or formal (or the same). These attributes correspond to the suspected interpersonal effects of speech characteristics such as speech melody. We receive information on the social and attitudinal effects of non-native intonation.

3. Results

The results show overall significantly different evaluations of the robots for the German participants ($N=45$; $\chi^2=15.5$; $p = 0.0165$),

near-significant differences for the native speakers of English ($N=18$; $\chi^2=10.75$; $p=0.0964$), yet no significant differences for the Danish participants ($N=52$; $\chi^2=3.24$; $p=0.777$). Post-hoc analyses show that Germans find the robot with Danish intonation dominant, while the Danish participants tend to perceive the robot with German intonation as formal ($p=0.06$).



Figure 2: The two Keepon robots taking turns asking questions

4. Discussion

The experiment illustrates the use of robots for the controlled investigation of the effects of linguistic features on the perception of the respective speaker and thus on language attitudes. In the current experiment, we used real speakers, whose utterances we modified using speech manipulation software; an alternative is to synthesize speech directly in order to eliminate all potential impacts of voice quality and speaker personality that may influence the attitude participants may hold towards speakers and/or the linguistic varieties they speak.

ACKNOWLEDGMENTS

This research was supported by the Danish Council for Independent Research under grant number 4180-00359A.

REFERENCES

- [1] Andrist, S., Spannan, E., and Mutlu, B. (2013). Rhetorical Robots: Making Robots More Effective Speakers Using Linguistic Cues of Expertise. In Proceedings of the 8th ACM/IEEE International Conference on Human Robot Interaction (HRI '13). IEEE Press, Piscataway, NJ, USA. 341-348.
- [2] Andrist, S., Ziadee, M., Boukaram, H., Mutlu, B., and Sakr, M. (2015). Effects of Culture on the Credibility of Robot Speech: A Comparison between English and Arabic. Proceedings of the Tenth ACM/IEEE International Conference on Human-Robot Interaction (HRI '15). ACM, New York, NY, USA. 157-164.
- [3] Boersma, Paul (2001). PRAAT: A system for doing phonetics by computer. In: Glot International 5(9/10), 341-347.
- [4] Bolinger, D., & Bolinger, D. L. M. (1989). *Intonation and its uses: Melody in grammar and discourse*. Stanford University Press.
- [5] Fischer, K. (2016). Robots as confederates: How robots can and should support research in the humanities. In *Proceedings of the Robophilosophy 2016 Conference, Aarhus, Denmark*, 2016.
- [6] Fischer, K., Niebuhr, O., Jensen, L.C. and Bodenhausen, L. (2019). Speech Melody Matters – How robots can profit from using charismatic speech. ACM Transactions in Human-Robot Interaction 9, 1, Article 4.
- [7] Giles, H., Wilson, P., & Conway, A. (1981). Accent and lexical diversity as determinants of impression formation and perceived employment suitability. *Language sciences*, 3(1), 91-103.
- [8] Grønnum, N. (2007). *Rødgrød med fløde: En lille bog om dansk fonetik*. København: Akademisk Forlag.
- [9] Lambert, W. E., R. Hodgson, R. C. Gardner and S. Fillenbaum. 1960. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60(1): 44–51.
- [10] Niebuhr, O. (2015). Gender Differences in the Prosody of German Questions. International Congress of Phonetic Sciences, Glasgow, UK, 1-5.
- [11] Soukup, B. (2013). On matching speaker (dis) guises—revisiting a methodological tradition. *Language (De) standardisation in Late Modern Europe: Experimental Studies*, 267-285.