On the sociolects of robots

MICHAEL PUCHER, Acoustics Research Institute (ARI), Austrian Academy of Sciences (ÖAW), Austria

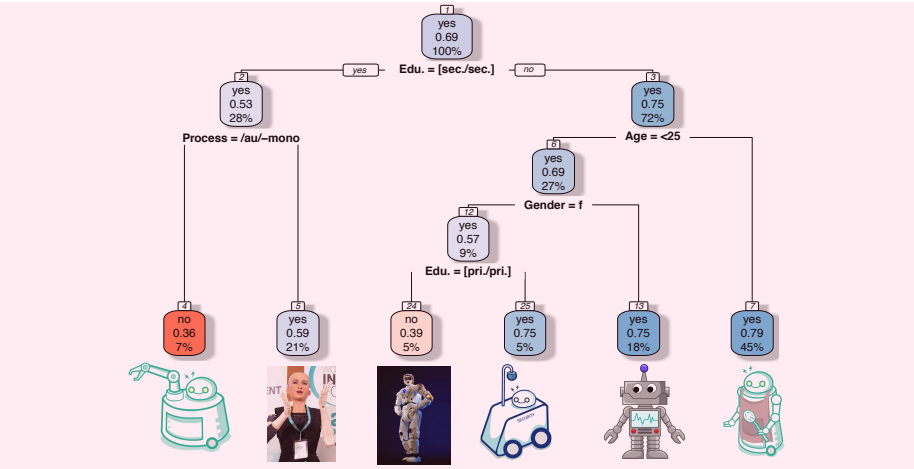

Fig. 1. Classification of Viennese sociolects associated with robot types (Robot images from Wikimedia).

In this text I want to approach the topic of artificial identity from a social perspective, asking the question if a robot can have a sociolect, how it would sound like, and if it is something that we should try to avoid. A sociolect is a language variation that is related to a speakers social status (education, age, gender). A robot can have a sociolect in two very different ways. The first possibility is the near-term human-robot sociolect that emerges simply from the fact that a robot has an interface based on human speech, and that no realization of human speech is free of social meaning. In the second, more science-fiction sense, a robot-robot sociolect could emerge through robots perceiving and evaluating speech from other robots to differentiate themselves. This would be a genuine robot-robot sociolect.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**.

Additional Key Words and Phrases: robots, sociolects

.

1 INTRODUCTION

Robots living in socially stratified human-robot and robot-robot worlds have been described in contemporary literature [4]. These environments lead to the emergence of socially stratified spoken language systems between humans and robots, and robots and robots. In analogy to linguistic research on human languages I want to term these spoken languages human-robot sociolects and robot-robot sociolects.

Manuscript submitted to ACM

Human-robot sociolects arise through the usage of human languages in human-robot interaction. Since human languages encode socially relevant information and stratification, these features are transferred to the robots in human-robot interaction. A robot using a male voice inherits features of male voices as evaluated by human speakers and listeners. A robot using a standard voice inherits the prestige of the standard, while a robot using a dialect voice inherits the lower prestige features of the dialect [6]. This has been described very well for spoken dialog systems in a recent UN report [2]. Although human-robot sociolects are based on features of sociolects from human languages they give rise to new research questions and directions since the theoretical possibility of adaptation to human listeners is huge for robots, an adaptation that a human speaker can only realize partially. That such an adaptation is also applied by human speakers can be seen in the different waves of sociolinguistics from variationist, to ethnographic and stylistic approaches that more and more emphasize the possibility of variation by the speaker [3].

There are two important features of sociolects, namely (i) variation and (ii) stratification. For the investigation of robot-robot sociolects these two features can be kept apart. It is possible that variation in robot speech occurs without any stratification. In human languages a sociolect is often related to a stratification, i.e. hierarchy between language varieties, so along these lines sociolects between robots could emerge, when one robot has an older speech interface with less capabilities, and a newer robot with a better interface is able to recognize this difference, and implicitly or explicitly uses this difference to distinguish itself from the other robot. In such a way a sociolect between robots would have emerged, based on technical progress. However, such a language game between robots only seems to make sense under the presence of a third human listener, such that these sociolects should be termed robot-robot-human sociolects. This would then be a genuine sociolect of robots.

Human-robot and robot-robot (robot-robot-human) sociolects can be located on all linguistic levels (sub-segmental, segmental, supra-segmental, lexical, syntactic, semantic, pragmatic) similarly to human sociolects. When related to specific speech modules used by a robot the arising variation is dependent on variation in human speech, and variation that relates to technical deficiencies or speech capabilities that outperform human capabilities.

Concerning the investigation of these robot sociolects methodical problems that arise with human sociolects such as the observer paradox could be overcome. This paradoxical situation arises when we want to "...find out how people talk when they are not being systematically observed; yet we can only obtain these data by systematic observation" [5]. Robots could observe themselves in an objective way and thus there would be no necessity for (i) the presence of someone in the role of the fieldworker, (ii) the presence of the recording device, (iii) or a specific task related to the investigation [3].

I have tried to sketch an answer to the question if robots can have sociolects and how they would sound like. An answer that remains to be given is if they should have sociolects, i.e. should live in a socially stratified world of robots. Even if we follow Asimov's laws such a stratification could easily arise among robots since according to Asimov's third law "A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws". Robots living in a socially stratified world would probably develop or be developed in a way that mirrors the hierarchies that are present in this world. This would include the development of social distinction to protect their own status. To avoid this, it seems to be necessary to overcome the social stratification of our world, where linguistic styles are used in an ideological way to force the recognition of the given order as natural [1].

2 BIOGRAPHY

Michael Pucher is a Senior Research Scientist at the Acoustics Research Institute (ARI) of the Austrian Academy of Sciences (ÖAW). He obtained his doctoral degree in Electrical and Information Engineering from Graz University of

Technology in 2007. In 2017 he received the venia docendi in Speech Communication at Graz University of Technology with a habilitation thesis on Speech Processing for Multimodal and Adaptive Systems. In the past he has worked on audio-visual speech synthesis, synthesis of singing, persona design for speech based systems, multimodal and spoken dialog systems, and semantic language modeling. He also made significant contributions in the area of speaker verification spoofing, where he showed how adaptive synthesizers can spoof a speaker verification system. Currently he is working on speech synthesis for language varieties, synthesis of animal vocalisations, digital phonetics, and sociophonetics. He has authored and co-authored more than 70 refereed papers in international conferences and journals. In 2010 he was involved in the commercial development of Leopold, the first synthetic voice for Austrian German. A list of publications and a detailed CV can be found on https://sociolectix.org.

REFERENCES

[1] Pierre Bourdieu. 1991. *Language and Symbolic Power.* Harvard University Press.

[2] EQUALS Skills Coalition. 2019. *I'd blush if I could: closing gender divides in digital skills through education.* https://unesdoc.unesco.org/notice?id=p::usmarcdef_0000367416

[3] P. Eckert. 2012. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. *Annu. Rev. Anthropol.* 41 (2012), 87–100.

[4] Kazuo Ishiguro. 2021. *Klara and the Sun.* Faber & Faber, London.

[5] William Labov. 1972. *Sociolinguistic Patterns.* University of Pennsylvania Press.

[6] Michael Pucher, Gudrun Schuchmann, and Peter Fröhlich. 2009. Regionalized Text-to-Speech Systems: Persona Design and Application Scenarios. In *Multimodal Signals: Cognitive and Algorithmic Issues*, Anna Esposito, Amir Hussain, Maria Marinaro, and Raffaele Martone (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 216–222.