Fall Semester 2022

Research Project: AVSR robot

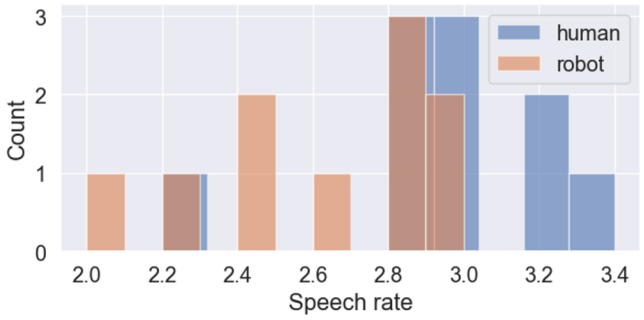
Speech characteristics when interacting with lip-reading social robot

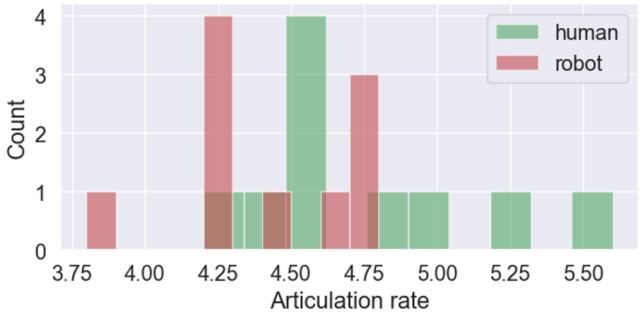
Estelle Chabanel

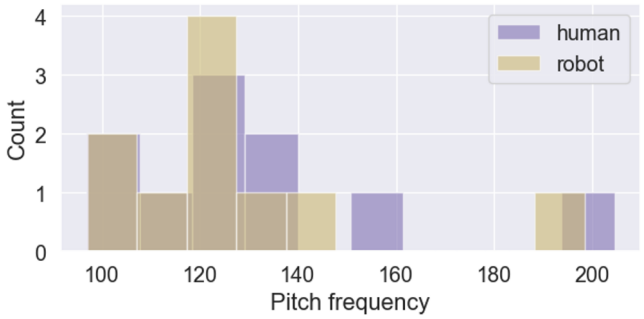
Supervisor: Daniel Tozadore

MOTIVATION

Understanding human robot interaction (HRI) is key to ensure social robots' success in educational matters. As studies on people's manner of speaking were found missing and of importance for the development of lip reading robots, this experiment explored speech characteristics in HRI.







METHODS

10 participants were video recorded in 2 conversations: one with a human and one with a robot with lip reading ability. In both conversations, 5 similar questions related to the 5 same topics were asked by the interlocutor. The videos were then used in a Visual Speech Recognition model and audios were analyzed to extract 3 speech characteristics: speech rate, articulation rate and mean pitch frequency.

RESULTS

As expected some differences in manner of speaking occur when comparing human conversation to HRI. Quantities of information delivered were smaller when addressing robots. The speech rate is also lower with robots, revealing a difference of perception of the interlocutor. Surprisingly, the articulation rate is higher with humans than robots, which could reveal a borness of the participants in the robot conversation. However, no important differences can be observed in the pitch frequency distribution: humans do not meaningfully change tone when addressing robots. Finally, these differences did not seem to impact the performances of the VSR robot, these ones depending on a lot of different factors.
The experiment should be realized with more participants to extract more reliable results and could be improved by making the robot interactions more natural.

Link to the code repo and report: https://github.com/EstelleChabanel/HRISpeechExperiment