Reversing Voice-Related Biases Through Haptic Reinforcement

Feras Al Taha
Centre for Intelligent Machines
McGill University
Montreal, Canada
feras.altaha@mail.mcgill.ca

Pascal E. Fortin
Centre for Intelligent Machines
McGill University
Montreal, Canada
pe.fortin@mail.mcgill.ca

Antoine Weill–Duflos
Centre for Intelligent Machines
McGill University
Montreal, Canada
antoinew@cim.mcgill.ca

Jeremy Cooperstock
Centre for Intelligent Machines
McGill University
Montreal, Canada
jer@cim.mcgill.ca

ABSTRACT
Biased perceptions of others are known to negatively influence the outcomes of social and professional interactions in many regards. Theses biases can be informed by a multitude of non-verbal cues such as voice pitch and voice volume. This project explores how haptic effects, generated from speech, could attenuate listeners' perceived voice-related biases formed from a speaker's voice pitch. Promising preliminary results collected during a decision-making task suggest that the speech to haptic mapping and vibration delivery mechanism employed does attenuate voice-related biases. Accordingly, it is anticipated that such a system could be introduced in the workplace to equalize people's contribution opportunities and to create a more inclusive environment by reversing voice-related biases.

Author Keywords
Speech; Haptics; Social Computing; Affective Computing

INTRODUCTION
Prior work has shown that people perceived as being more dominant tend to take more speaking time. This can reduce effectiveness of a group meeting and result in a less preferable outcome than a meeting with equal speaking opportunities [4]. In addition, Takayama et al. has shown that an individual's influence on the decision-making process is correlated with how others perceive that individual's dominance [11].

UIST'18 Adjunct, October 14–17, 2018, Berlin, Germany.

ACM ISBN 978-1-4503-5949-8/18/10.
http://dx.doi.org/10.1145/3266037.3266101

While the speaker's physical height [13], mean voice pitch, pitch variability, mean loudness, mean resonance and resonance variability were all shown to influence perceived properties from the speaker [7, 5, 2, 8, 12, 6], this work focuses specifically on voice-pitch-related bias and its influence on listeners' decision-making. Many studies demonstrated, within and between genders, that a higher voice pitch is generally associated with less physical dominance [5, 2, 8, 12] and less leadership abilities [6] than its lower pitch counterpart.

Attempts have been made to manipulate perception of speech by altering speakers' vocal properties. An example of such a system by Aucouturier et al. covertly modified the emotional tone of participants' voices toward happiness, sadness, or fear, in real time while they talked [1]. Similarly, Rachman et al. created an open-source software platform to manipulate speech and emotional cues by modifying pitch and other voice characteristics [9]. While these systems successfully altered the perceived properties of the speakers, we argue that modifying users' voices is neither practical nor suitable for attenuating voice biases as it changes vocal identity. Making every speaker perceived equally by modifying their voice would imply making everyone sound the same.

This project introduces a speech-to-haptics processing pipeline that allows the exploration of the influence of haptics on speech perception without altering speakers' voices. Promising preliminary results are presented and outline the potential of the proposed system at attenuating voice-related biases in binary decision-making processes, a scenario that can frequently be encountered in professional and social meetings.

PROPOSED SYSTEM
A system was built to generate haptic effects from speech in real time with the objective of modifying the perceived dominance of speakers without altering their vocal properties. Less dominant voices would be perceived as more intense

vibrations than naturally dominant voices all while keeping the vibrations in the background and non-distracting.

The speech signal is first acquired using a microphone. The system extracts the fundamental frequency of the voice, and uses it to modulate the gain of the captured signal. A higher pitched voice is mapped to larger gain, while a lower pitched voice will result in an attenuation of the signal. Then, the signal's frequency spectrum is shifted to a range that is more tactually perceptible (50 to 300 Hz) [3, 10]. Finally, the signal is low-pass filtered to remove higher frequency harmonics that would be audible, before being amplified and presented using voice coil actuators attached under the participant's office chair. The result is a haptic signal nearly perfectly synchronized to speech that is anticipated to modulate perceived dominance of speakers through manipulation of the vibration intensity.

PILOT STUDY

A preliminary pilot study was designed to determine if the proposed system and speech-to-haptic mapping can indeed modify perceived dominance, influence listeners' decision-making and attenuate voice-pitch-related biases. Pairs of text-to-speech generated sentences, instructing to chose one of two options, were presented to participants with and without the proposed haptic reinforcement. For each pair of sentences, a higher and lower pitch version of the same voice was used (10 semitone difference). Participants were asked to select one of the two options by pressing the keyboard key corresponding to the spoken instruction of their choosing.

Results

Seven members of the authors' research group, with no prior knowledge of the project, agreed to participate in the pilot study.

As shown in Figure 1, preliminary results suggest that the introduction of haptics does influence subjects in choosing the higher-pitched (55% of choices) over the lower-pitched instruction (45% of choices) more frequently. A larger variance can also be observed, which we argue could be caused by the haptic effects being too invasive and insufficiently subtle, making some participants uncomfortable and creating an avoidance of the haptically stronger instruction. This can be avoided by tuning the system to produce less intrusive and more subtle vibrations.

However, unlike what was advocated by prior literature [5, 2, 6], participants did not seem to follow the lower pitch instructions more than the higher pitch instructions in the no-haptics condition. We hypothesize that this could be due to the usage of a female voice and an overly open question. The choice being left free, it is possible that the participants made their decisions under different criteria such as attractiveness or dominance. In such case, using a female voice could provide opposed results because high pitched voices are perceived as more attractive up to a certain point [5, 2]. Additionally, despite trying to pick neutral instruction sentences, some of the instructions might have had a bias that made participants chose one option over the other, regardless of pitch.

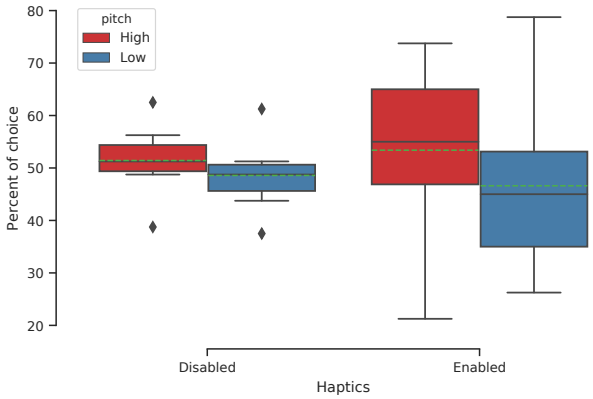

Figure 1. Results of pilot experiment

CONCLUSION

Results from this pilot study demonstrated that haptic enhancement of the voice does have a biasing effect on which instructions participants would be following. Additionally, the undesired effect of avoiding stronger haptic effects could be interesting to investigate and is a clue for future works. If the biasing effect can be controlled, we could achieve a system that equalizes all voices and neutralize voice-related biases such that speaking opportunities will be equal for all and unaffected by voice biases. Such a system could make workplaces more inclusive to everyone and improve meeting outcomes.

REFERENCES

1. Jean-Julien Aucouturier, Petter Johansson, Lars Hall, Rodrigo Segnini, Lolita Mercadié, and Katsumi Watanabe. 2016. Covert digital manipulation of vocal emotion alter speakers' emotional states in a congruent direction. 113, 4 (2016), 948–953. DOI: http://dx.doi.org/10.1073/pnas.1506552113
2. Barbara Borkowska and Boguslaw Pawlowski. 2011. Female voice frequency in the context of dominance and attractiveness perception. 82, 1 (2011), 55–59. DOI: http://dx.doi.org/10.1016/j.anbehav.2011.03.024
3. Vincent Hayward and Oliver R. Astley. 1996. Performance measures for haptic interfaces. In *Robotics research*. Springer, 195–206.
4. Hayley Hung, Yan Huang, Gerald Friedland, and Daniel Gatica-Perez. 2011. Estimating dominance in multi-party meetings using speaker diarization. 19, 4 (2011), 847–860.
5. Benedict C. Jones, David R. Feinberg, Lisa M. DeBruine, Anthony C. Little, and Jovana Vukovic. 2010. A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. 79, 1 (2010), 57–62. DOI: http://dx.doi.org/10.1016/j.anbehav.2009.10.003
6. Casey A. Klofstad, Rindy C. Anderson, and Susan Peters. 2012. Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. (2012), rspb20120311. DOI: http://dx.doi.org/10.1098/rspb.2012.0311

7. Sei Jin Ko, Melody S. Sadler, and Adam D. Galinsky. 2015. The Sound of Power: Conveying and Detecting Hierarchical Rank Through Voice. 26, 1 (2015), 3–14. DOI:http://dx.doi.org/10.1177/0956797614553009
8. David Andrew Puts, Carolyn R. Hodges, Rodrigo A. Cárdenas, and Steven J. C. Gaulin. 2007. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. 28, 5 (2007), 340–344. DOI: http://dx.doi.org/10.1016/j.evolhumbehav.2007.05.002
9. Laura Rachman, Marco Liuni, Pablo Arias, Andreas Lind, Petter Johansson, Lars Hall, Daniel Richardson, Katsumi Watanabe, Stéphanie Dubal, and Jean-Julien Aucouturier. 2018. DAVID: An open-source platform for real-time transformation of infra-segmental emotional cues in running speech. 50, 1 (2018), 323–343. DOI: http://dx.doi.org/10.3758/s13428-017-0873-y
10. F. Reynier and V. Hayward. 1993. Summary of the Kinesthetic and Tactile Function of the Human Upper Extremities. *McGill Research Center for Intelligent Machines, McGill University, Montreal, Canada* (1993).
11. Leila Takayama and Clifford Nass. 2010. Throwing voices: The psychological impact of the spatial height of projected voices. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (2010). ACM, 91–94.
12. Kyle James Tusing and James Price Dillard. 2000. The Sounds of DominanceVocal Precursors of Perceived Dominance During Interpersonal Influence. 26, 1 (2000), 148–172. DOI: http://dx.doi.org/10.1111/j.1468-2958.2000.tb00754.x
13. Christopher D. Watkins, Paul J. Fraccaro, Finlay G. Smith, Jovana Vukovic, David R. Feinberg, Lisa M. DeBruine, and Benedict C. Jones. 2010. Taller men are less sensitive to cues of dominance in other men. 21, 5 (2010), 943–947. DOI: http://dx.doi.org/10.1093/beheco/arq091