OWN-VOICE PITCH PERCEPTION ISN'T JUST ABOUT BONE CONDUCTION

Maddy Walter, Sydney Norris, Grace Bengtson, Liang Kai Fong, Bryan Gick
Department of Linguistics, University of British Columbia, Vancouver, Canada

1 Introduction

Awareness of one's own voice starts during infancy [2], and own-voice egotism is present later in life, i.e., individuals prefer voices like their own and rate their voice as more attractive than others do [5; 11]. However, dissatisfaction with hearing one's recorded voice [4] suggests possible misjudgement of own-voice characteristics.

Own-voice perception differences are often attributed to air vs. bone conduction. Studies have challenged the notion that air-conduction differences completely account for this perceptual contrast [6; 7], describing own-voice as a "multi-modal construct" [9, pg. 1]. However, previous methodology does not account for own-voice egotism, and the complex influence of social expectations on perceptual experience remains uncertain. We addressed this research gap by testing how socially-based, gendered pitch expectations [12] impact own-voice perception, while controlling for own-voice egotism.

The existing literature supports two predictions for how participants could perceive the pitch of their recorded voice. A *Strong Bone Conduction Hypothesis* [13] suggests that live-spoken voices are heard as if through a low-pass filter, so that one's own recorded voice sounds comparatively higher in pitch (Figure 1).

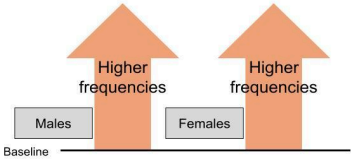

Figure 1: Schematized recorded own-voice perception according to a *Strong Bone Conduction Hypothesis*.

Alternatively, a *Gendered Pitch Expectation Hypothesis* acknowledges socially-based gendered expectations for pitch, and considers own-voice egotism. This hypothesis predicts that if individuals believe that their recorded voice sounds inauthentic, female-identifying participants will believe their recorded voice has been lowered, while male-identifying participants will believe theirs has been raised (Figure 2).

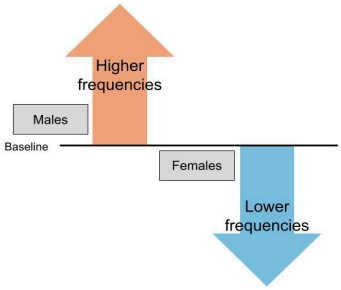

Figure 2: Schematized recorded own-voice perception according to a *Gendered Pitch Expectations Hypothesis*.

2 Methods/Analysis

101 participants were recruited via UBC's Linguistics Outside the Classroom (LOC) SONA System. Strict exclusionary criteria were applied in order to eliminate possible confounding factors. Participants were between the ages of 18 and 50 and had no history of hearing or laryngeal impairments. Those who had received professional voice training or speech therapy were excluded. Participants were also excluded if they had lived in North America for less than ten years, as gender-based pitch expectations may be culturally defined [14]. Participants who did not identify as male or female were also excluded at this stage to avoid the assumption of pitch expectations with less-studied gender identities. Finally, participants were excluded if they had experienced any cold or flu-like symptoms in the previous week, as these symptoms often affect vocal quality [1]. The remaining data consisted of 31 female-identifying and 15 male-identifying participants, totalling 46 observations measured.

Participants were given privacy to answer survey questions about their gender identity and their sense of belonging in that social group. While accompanied by a researcher, participants were then recorded in a sound booth saying the neutral statement: "Some birds build nests in tall trees. Some mice burrow in the ground." Audio was recorded on a dynamic microphone using Audacity. After leaving the sound booth for five to ten minutes, participants returned and were informed that their recording may have been pitch-altered. To ensure a baseline of understanding before hearing their own recording, participants were played example recordings of a voice that had been raised and lowered. Next, hearing the unaltered recording of their voice once, participants were asked to rate its authenticity and how to adjust the pitch to sound like themselves.

Participants rated the authenticity of their voices on a 4-point scale, with options to respond that their voice was "Not altered", "Altered slightly", "Altered substantially", or "That was not a recording of my voice". In this analysis, "Altered slightly" and "Altered substantially" have been grouped. Participants then rated how their voice was pitch-adjusted on a 7-point scale, with options ranging from "Substantially lowered/raised," "Moderately lowered/raised," "Slightly lowered/raised," or "Unaltered."

Participant responses were visualized in Python.

3 Results

While all participants believed that the recording was of their own voice, the majority of participants (87.1% of females and 86.7% of males) believed that their recorded voice had been altered. Figure 3 plots the proportion of

participants who rated the recording as authentic or 'unaltered'.

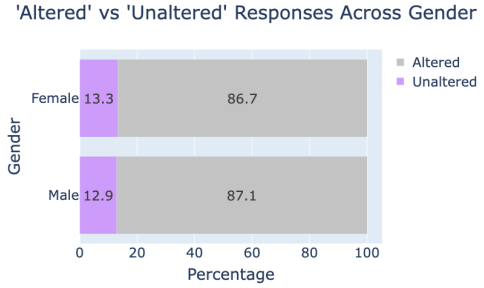

Figure 3: Frequency of 'unaltered' responses across genders.

Among participants who believed their recorded voice had been altered, a majority in both gender groups reported that their voice had been lowered rather than raised (females: 54.8% vs. 32.3%; males: 53.3% vs. 33.3%). While the difference across gender groups did not reach statistical significance ($\chi^2(1, N = 46) = 2.57$, $p = 0.11$), the difference between raised and lowered responses did ($\chi^2(1, N = 46) = 33.93$, $p = 5.714e\text{-}09$), suggesting a robust bias across genders. Figure 4 plots the frequency of each response on the pitch-adjustment scale across both genders.

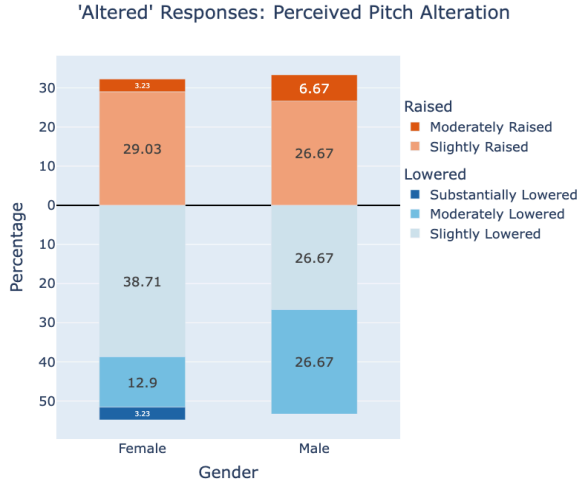

Figure 4: Frequency of response type across genders among participants who believed their voice was altered.

4 Discussion

Neither the *Strong Bone Conduction Hypothesis* nor the *Gendered Pitch Expectation Hypothesis* was supported, as both male- and female-identifying participants thought that their recorded voice had been lowered (Figure 5).

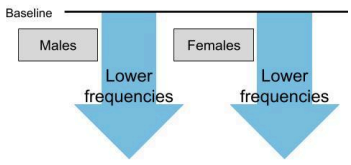

Figure 5: Both gender groups believed their recorded voices were lowered, counter to both hypotheses.

This study takes a step towards a fuller picture of own-voice perception. Crucially, these findings implicate factors beyond air conduction differences that contribute to perceptual misjudgement of recorded own-voice, as perception of lowered frequencies counters a *Strong Bone Conduction Hypothesis*. They also suggest that individuals are not judging own-voice characteristics accurately regardless of gender identity. Most individuals believed their voices had been altered, failing to meet their own pitch expectations. A robust unidirectional bias across genders counters a *Gendered Pitch Expectation Hypothesis*.

These results help bridge the gap between conflicting findings of own-voice egotism [5; 11] and dissatisfaction with hearing one's own recorded voice [4], highlighting the complexity of own-voice perception as a multidimensional phenomenon.

What is really going on in own-voice perception? We consider here a third hypothesis: a *Strong Perceptual Hypothesis (SPH)*. Contrary to the *Strong Bone Conduction Hypothesis*, which is a purely sensory hypothesis, the SPH holds that our perceptual systems account for sensory effects. This predicts an opposite effect from what we expect, with over-compensation when participants believe their recording has been altered. Similar effects have been observed in temporal windows of sensory integration: as suggested by [3] and [8], our perceptual systems anticipate purely sensory effects in auditory-tactile and auditory-visual coupling, respectively. The present results indicate the same may be true of bone conduction and pitch, supporting the view that, while a *Strong Bone Conduction Hypothesis* may correctly describe a sensory effect, the perceptual system's response is better described by a *Strong Perceptual Hypothesis*.

Acknowledgments

This project was funded by an NSERC Discovery grant. We would like to acknowledge the contributions to data collection from Shannon Colcleugh.

References

[1] Bové, M. J., Kansal, S., & Rosen, C. A. (2008). Influenza and the vocal performer: Update on prevention and treatment. *J. Voice, 22*(3), 326-332.

[2] Dondi, M., Simion, F., & Caltran, G. (1999). Can newborns discriminate between their own cry and the cry of another newborn infant?. *Dev. Psych., 35*(2), 418-426.

[3] Gick, B., Ikegami, Y., & Derrick, D. (2010). The temporal window of audio-tactile integration in speech perception. *The Journal of the Acoustical Society of America, 128*(5), EL342-EL346. https://doi.org/10.1121/1.3505759

[4] Holzman, P. S., & Rousey, C. (1966). The voice as a percept. *J. Pers. & Soc. Psych., 4*(1), 79-86.

[5] Hughes, S. M., & Harrison, M. A. (2013). I like My Voice Better: Self-Enhancement Bias in Perceptions of Voice Attractiveness. *Perception, 42*(9), 941-949.

[6] Kimura, M., & Yotsumoto, Y. (2018). Auditory traits of "own voice". *PloS One, 13*(6), e0199443-e0199443.

[7] Maurer, D., & Landis, T. (1990). Role of bone conduction in the self-perception of speech. *Folia Phon. et Logo., 42*(5), 226-229.

[8] Munhall, K. G., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Perception & Psychophysics, 58*(3), 351- 362. https://doi.org/10.3758/BF03206811

[9] Orepic, P., Kannape, O. A., Faivre, N., & Blanke, O. (2023). Bone conduction facilitates self-other voice discrimination. *Roy. Soc. Open Sci., 10*(2), 221561-221561.

[10] R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R- project.org/.

[11] Peng, Z., Wang, Y., Meng, L., Liu, H., & Hu, Z. (2019). One's own and similar voices are more attractive than other voices. *Aus. J. Psych., 71*(3), 212-222.

[12] Peng, Z., Hu, Z., Wang, X., Jiao, T., Li, H., & Liu, H. (2021). Gender context modulation on the self‐enhancement effect of vocal attractiveness evaluation. *PsyCh J., 10*(6), 858-867.

[13] Tonndorf, J. (1976). Bone conduction. In *Auditory System. Handbook of Sensory Physiology (eds. Keidel W.D., Neff W.D.), 37-84*. Berlin, Germany: Springer.

[14] Zimman, L., & Yuasa, I. P. (2011). Culture and gender of voice pitch: A sociophonetic comparison of the Japanese and Americans. *Lang. in Soc., 40*(2), 251-252.