Undergraduate Thesis Prospectus

Evidence of Bias in Voice over Internet Protocol Audio Compression

(technical research project in Systems Engineering)

Disrupted Voices: The Impact of Inequitable Constraints in Digital Communication Technologies Affecting Sex Representation

(Sociotechnical research project)

by

Elizabeth Recktenwald

November 8, 2024

Technical project collaborators:

Lucas Vallarino

Catherine Nguyen

Maddie Sullivan

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Elizabeth Recktenwald

Technical advisor: Dr. Matthew Bolten, Department of Systems Engineering

STS advisor: Peter Norton, Department of Engineering and Society

General Research Problem

Bias in Digital Voice Communication Systems Affecting High-Frequency Voices

How do biases in digital voice communication systems disadvantage high-frequency voices, particularly those of individuals identifying as female?

Digital voice communication systems, including Voice over Internet Protocol (VoIP) and voice recognition technologies, are central to modern communication but often exhibit biases that disproportionately affect high-frequency voices, typically women's. These voices may experience distortion, reduced intelligibility, and lower quality, leading to challenges in clarity and equality within digital interactions (Bolton, 2023).

This issue holds both technical and social significance. Technically, current system designs struggle to accurately process high-frequency voices, leading to degraded performance and diminished reliability. Socially, these biases limit fair representation in domains reliant on clear voice transmission, such as business and emergency response. Addressing these biases is crucial for advancing the inclusivity and functionality of global communication platforms.

Evidence of Bias in Voice over Internet Protocol Audio Compression

How can we study the accuracy of digital audio codecs to minimize frequency distortion in high-frequency voices?

This capstone project, supervised by Dr. Matthew Bolten in the Systems Engineering department, explores how digital audio codecs manage high-frequency voices in VoIP communication. Project team members include Lucas Vallarino, Catherine Nguyen, and Maddie Sullivan.

Digital voice communication technologies like VoIP and AI-driven voice recognition systems are essential in today's communication landscape but frequently exhibit biases that impact high-frequency voices. Moore and Tan (2003) note significant degradation in speech quality with narrowed bandwidths, especially for typical telephone ranges (313 to 3547 Hz), resulting in distortion and reduced intelligibility for high-pitched voices. These issues compromise clarity and contribute to gender-based disparities in digital communication.

Research by Lee (2024) and Fernández Gallardo et al. (2018) has shown that biases in codec designs limit user access to reliable voice transmission, with dialect and gender-related factors further exacerbating voice clarity issues. AI models can even reinforce stereotypes by associating dialectal variations with lower-status roles (Lee, 2024). Addressing codec limitations in handling high-frequency voices could foster a more equitable digital communication environment.

Existing codecs, especially those optimized for low bit rate, struggle to maintain high-frequency voice fidelity, particularly in lower-quality settings where frequency distortion is most pronounced. Dr. Bolton's (2023) study indicates that high-frequency voices experience disproportionate intelligibility loss, intensifying gender-related disparities in voice transmission. The concept mapping of a codec is depicted in the figure below to promote understanding.

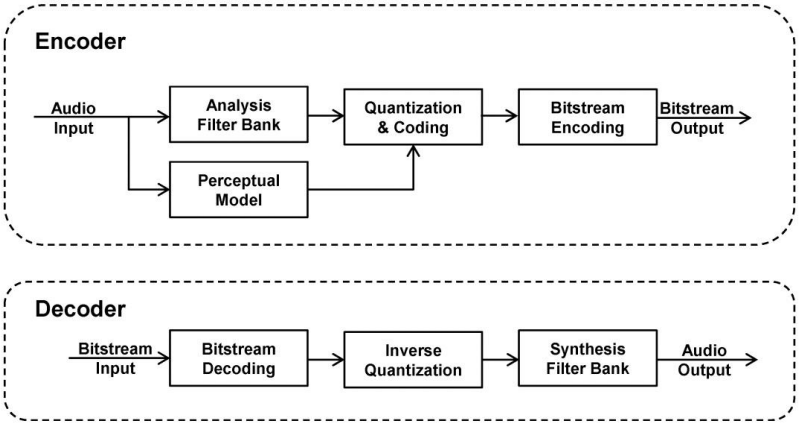

Figure 1: Schematic principle of a codec (Herre and Dick 2019, p.2)

The project has two main goals: to identify codec configurations that minimize frequency distortion and to recommend improvements that enhance audio fidelity for high-frequency voices, focusing on inclusivity for female users.

Unusual constraints in this research include restricted access to proprietary codecs and the need for specialized testing equipment, which may limit the range of codecs for in-depth analysis.

The current state of the art in VoIP technology has advanced to facilitate efficient voice transmission, yet it fails to adequately represent diverse voice types, often leading to "unintended distortions that disproportionately affect women and gender-diverse individuals" (Gallardo & Sanchez-Iborra, 2019). While existing systems prioritize efficiency, there is a need for innovations that support more inclusive voice quality and equitable representation (AJL, n.d.).

To evaluate codec performance, the team will apply a mixed-methods approach. Objective measures like Perceptual Evaluation of Speech Quality (PESQ) and Perceptual Evaluation of Audio Quality (PEAQ) will assess intelligibility, while subjective listening tests in MATLAB will gauge perceived audio quality. Comparative analyses across codec configurations will identify settings that best preserve high-frequency voice clarity.

The completed project will provide an analysis of codec performance on high-frequency voices, particularly AMR, OPUS, and CODEC2, and will recommend configurations to improve clarity. These findings aim to enhance inclusivity and promote gender equity in digital communication spaces.

Disrupted Voices: The Impact of Inequitable Constraints in Digital Communication Technologies Affecting Sex Representation

How have equity advocacies identified and addressed biases in digital communication systems that disproportionately impact individuals with high-frequency voices?

Digital communication technologies have transformed interpersonal and professional interactions, yet these systems often introduce biases that disproportionately affect certain voice frequency ranges. This exclusion of particular voice types raises critical concerns about equitable representation in digital spaces. Exclusion from effective communication can undermine both access and inclusivity, potentially reinforcing barriers to information and professional opportunities for those impacted by these biases.

Research has established foundational evidence of bias within digital systems. S Siegert and Niebuhr (2021) examine audio compression's detrimental effect on the perceived charisma of female voices in remote settings. "Prior studies have concluded that charisma perception of speakers plays an important role in how attentive listeners are, the impact on their motivation and how well they perform in numerous tasks (ibid.)." Siegert et al. (2021) further observe that "any signal degradation made the speakers' emotions less intelligible for listeners." Despite growing recognition of these issues, there remains a gap in addressing how low-band codecs impact communication clarity for high-frequency voices, thereby limiting inclusivity within digital environments.

A technical baseline for audio codec standards were established by Bosi and Goldberg (2012), whose work on digital audio coding emphasized optimal bit rate cutoffs—criteria that, while important for efficient data transmission, inadvertently form a basis for some of the previously observed biases in Bolton's (2023) work. These standards highlight an industry tendency to optimize for specific voice profiles, often at the expense of others, including those with high-frequency voices. These unintended biases show the need for targeted advocacy to address these technical constraints and promote equitable audio standards.

Several equity advocacies are working to counter these biases. Women in Technology (WIT) is one such organization, committed to advancing gender equity within the tech industry. "WIT began as a small group of women in the Washington DC tech community, united by the belief that our voices are stronger together," showcasing their dedication to amplifying women's representation in technology. Since its founding in 1994, WIT has provided networking and development resources to elevate women's professional impact, reinforcing their presence as a leader in advocacy for women's voices in tech.

Similarly, the Women's Audio Mission (WAM) supports the representation of women in audio and media by creating spaces where female perspectives in digital audio technology are celebrated and promoted. WAM shows that "fewer than 5% of the people creating the sounds, music, and media in the daily soundtrack of our lives are women or gender-expansive individuals," highlighting the gender disparity in creative tech fields. Through their award-winning curriculum, which merges art and STEM, WAM has inspired over 25,000 participants to amplify their voices as digital innovators.

UN Women also shows the need for equitable digital solutions that ensure fair treatment across all voice frequencies, reflecting the global importance of these issues. Through its advocacy, UN Women promotes research-based improvements to remove biases that affect

gender equity in digital spaces. Bahous's call for action— "we must ensure that technology amplifies women's voices so that they may have a stronger impact on the future course of our world"—resonates as a mission for ensuring that digital technologies serve as inclusive platforms (Bahous, 2023).

The Algorithmic Justice League (AJL) plays a significant role in addressing biases in artificial intelligence, which frequently extend into digital audio systems. AJL's mission is to "raise awareness about the impacts of AI, equip advocates with empirical research, build the voice and choice of the most impacted communities, and galvanize researchers, policymakers, and industry practitioners to mitigate AI harms and biases" (AJL, n.d.). Through initiatives like the #FreedomFlyers Campaign, AJL encourages individuals to take control of their biometric rights and holds agencies accountable for their use of AI technologies. Their work serves as a reminder of the far-reaching impacts of unchecked AI bias, including issues tied to voice representation in digital communication.

References

Admin, M. (2024). *WIT History*. Women in Technology. https://www.womenintechnology.org/

AJL (n.d.). Algorithmic Justice League. Unmasking AI harms and Biases.

Bahous, S. (2023, March 14). *Technology and gender equality-bringing women and girls to the centre of Innovation*. United Nations. https://www.un.org/en/un-chronicle/technology-and-gender-equality%E2%80%94bringing-women-and-girls-centre-innovation

Başkent, D., & Chatterjee, M. (2010). *Recognition of temporally interrupted and spectrally degraded sentences with additional unprocessed low-frequency speech*. Hearing Research, 270(1-2), 127–133.

Bolton, M. (2023). *Preliminary Evidence of Sexual Bias in Voice over Internet Protocol Audio Compression*. RDCU.be.

Bosi, M., & Goldberg, R. E. (2012). *Introduction to Digital Audio Coding and Standards*. Springer, Heidelberg.

Daum, A. (2023, August 14). *Impact of Audio Artifacts on Perceptions of Male and Female Voices*. Erlangen, Germany; Friedrich-Alexander-Universität Erlangen-Nürnberg Department of Computational Linguistics.

Gallardo, L. F., & Sanchez-Iborra, R. (2019, July 26). On the Impact of Voice Encoding and Transmission on the Predictions of Speaker Warmth and Attractiveness. *ACM Transactions on Knowledge Discover from Data 13* (4).

Herre, J., & Dick, S. (2019). Psychoacoustic models for perceptual audio coding—Tutorial review. *Applied Sciences 9* (14), 2854.

Lee, T. (2024, Sep. 17). Ai is biased against speakers of African American English, study finds. *University of Chicago News.*

WAM (2024, Jan. 8). Women's Audio Mission. About.

Moore, B.C., & Tan, C.T. (2003). Perceived naturalness of spectrally distorted speech and music. *Journal of the Acoustical Society of America,* 114(1), 408–19.

Siegert, I., & Niebuhr, O. (2021). Case report: Women, be aware that your vocal charisma can dwindle in remote meetings. *Frontiers in Communication 5.*

Voran, S. D., & Catellier, A. A. (n.d.). *Speech Codec Intelligibility Testing in Support of Mission-Critical Voice Applications for LTE*. NIST.