

Misheard and Misgendered? The Accuracy of Speaker Gender Identity Attribution

Katelyn Taylor and Kirsty McDougall
University of Cambridge, klt56|kem37@cam.ac.uk

Phonetic and sociolinguistic research has examined how speaker sex influences their speech patterns, alongside social factors like age, social class, and sexuality. However, previous studies on speaker identification and perception have predominantly focused on cisgender speakers, overlooking gender-diverse voices [1], [2], [3]. Importantly, while sex (e.g. male/female assignment at birth) has been widely studied in phonetics, less attention has been paid to whether listeners can identify a speaker's gender identity from their speech. The present study builds on a study that found a transgender speaker was most accurately identified in a voice parade (containing 7 cisgender and 1 transgender speaker), raising the question of whether listeners can accurately determine gender identity from speech [4].

Eighteen speakers of Southern Standard British English participated: 6 cisgender (3 male, 3 female), 6 non-binary [NB] (3 assigned male at birth [AMAB], 3 assigned female at birth [AFAB]), 6 transgender (3 transmasculine, 3 transfeminine). In a listening test, 61 listeners were asked to determine whether speech samples were produced by a cisgender, non-binary, or transgender speaker, to rate their confidence, to assess how masculine or feminine they perceived each voice to be, and to select relevant phonetic cues from nine predefined features (e.g. pitch, intonation, nasality, creakiness, breathiness) with an additional open-text option.

Following a pilot study, it was hypothesised that listeners would be unable to identify the exact gender identity of the non-binary and transgender speakers but would be able to classify the speakers as being cisgender versus gender queer somewhat accurately. Results showed that listeners were more successful in identifying the speakers as cisgender, as opposed to non-binary or transgender (Figure 1). Non-parametric statistical analyses (Wilcoxon and Kruskal-Wallis tests) confirmed that when the scoring system was broadened to create a "gender queer" category containing speakers who identify as either NB or transgender, listener accuracy increased for every NB (56% for NB AFAB, 13% for NB AMAB) or transgender speaker (63% for transfeminine, 22% for transmasculine). This suggests that listeners could broadly recognise gender queer voices but struggled to assign the specific gender identity accurately. Statistical analysis also revealed that the speaker's sex (whether they were AMAB or AFAB) did not play a significant role in listener accuracy.

This indicates that phonetic cues are likely to be relied on to distinguish gender queer speakers from cisgender ones. Further analysis of listeners' confidence and their perception of relevant phonetic characteristics will be presented.

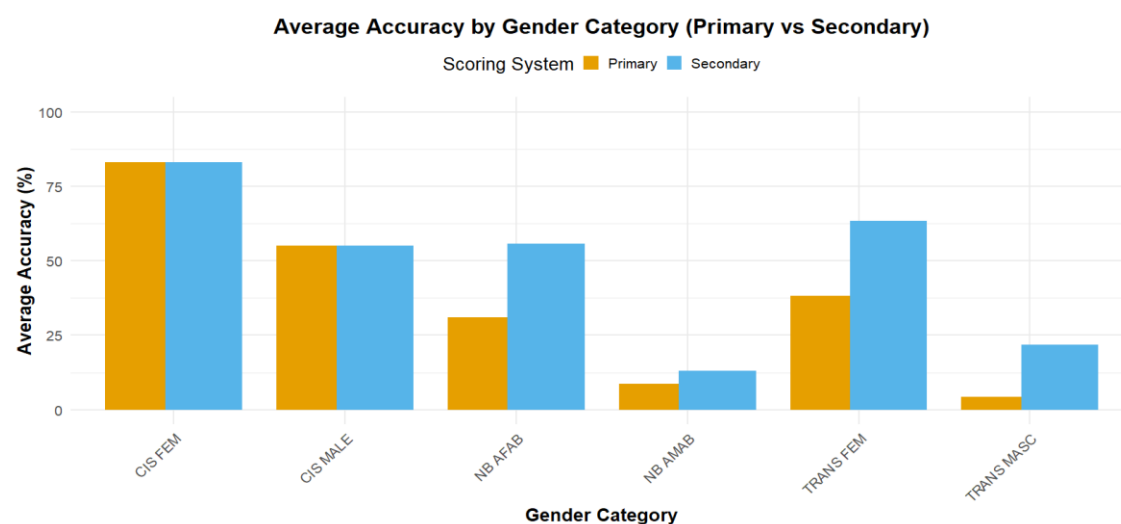


Figure 1: A graph showing how often speakers' gender identities were correctly attributed. The Primary scoring system refers to the exact correct gender identity being selected by the listener, and the secondary scoring system refers to speakers being marked as either cisgender or gender queer.

References

- [1] Zimman, L. (2018). Transgender voices: Insights on identity, embodiment, and the gender of the voice. *Language and Linguistics Compass*, 12(8), e12284.
- [2] Zimman, L. (2021). Gender diversity and the voice. In *The Routledge handbook of language, gender, and sexuality* (pp. 69-90). Routledge.
- [3] Merritt, B. (2023). Speech beyond the binary: Some acoustic-phonetic and auditory-perceptual characteristics of non-binary speakers. *JASA Express Letters*, 3(3).
- [4] Taylor, K., Gully, A., & Daffern, H. (2024). Familiar and Unfamiliar Speaker Identification in Speech and Singing. ISCA.