3-15-2023

Speech Beyond the Binary: Some Acoustic-Phonetic and Auditory-Perceptual Characteristics of Non-Binary Speakers

Brandon Merritt
bmmerritt@utep.edu

JASA EXPRESS LETTERS | ARTICLE | asa.scitation.org/journal/jel

Speech beyond the binary: Some acoustic-phonetic and auditory-perceptual characteristics of non-binary speakers

Brandon Merritt[a]
Department of Rehabilitation Sciences, The University of Texas at El Paso, El Paso, Texas 79902, USA
bmmerritt@utep.edu

Abstract: Speech acoustics research typically assumes speakers are men or women with speech characteristics associated with these two gender categories. Less work has assessed acoustic-phonetic characteristics of non-binary speakers. This study examined acoustic-phonetic features across adult cisgender (15 men and 15 women) and subgroups of transgender (15 non-binary, 7 transgender men, and 7 transgender women) speakers and relations among these features and perceptual ratings of gender identity and masculinity/femininity. Differing acoustic-phonetic features were predictive of confidence in speaker gender and masculinity/femininity across cisgender and transgender speakers. Non-binary speakers were perceptually rated within an intermediate range of cisgender women and all other groups.

[Editor: Douglas D. O'Shaughnessy] https://doi.org/10.1121/10.0017642

Received: 24 November 2022 **Accepted:** 1 March 2023 **Published Online:** 15 March 2023

1. Introduction

Per the American Psychological Association, individuals may be broadly considered as *cisgender* (i.e., a person whose gender identity and gender expression align with sex assigned at birth) or *transgender* (i.e., the full range of people whose gender identity and/or gender role do not conform to what is typically associated with their assigned sex at birth).[1] Transgender is an umbrella term that is inclusive of individuals who align with a binary gender categorization (e.g., transgender men and transgender women) and individuals whose gender identity is not positioned within a binary gender framework (e.g., non-binary). Studies on the psychosocial experiences of non-binary individuals indicate that they constitute a distinct identity apart from binary cisgender and transgender men and women.[2–4] For example, non-binary individuals are more likely than cisgender and transgender men and women to reject the belief that gender dictates individual personality and behaviors.[3,4] Non-binary individuals also tend to place less importance on "passing" as a particular gender.[2] They may therefore be less likely to modify their speech characteristics to align with those of cisgender men and women.

Emerging research indeed suggests that non-binary speakers may selectively adopt speech characteristics that are distinct from those of cisgender and transgender men and women and that communicate their non-binary identity to listeners.[5–8] For example, in a study of transmasculine and non-binary participants, speakers who identified as "straight men" had the lowest /s/ center of gravity across all speakers, while speakers who used labels such as "boy" or "genderqueer" had the highest.[8] Some non-binary speakers have also been found to combine a relatively low fundamental frequency (f_o) (a masculinizing perceptual feature) with increased use of high-rising terminals (a feminizing perceptual feature).[6] These findings suggest that a constellation of idiosyncratic (e.g., personal sense of gender identity and sexuality) and systemic (e.g., social experiences) factors may collectively inform non-binary speakers' linguistic *style*, where individual acoustic-phonetic features are combined to construct and convey particular social meaning.[9] Importantly, many of these features (and their stylistic variations) occur as sociocultural phenomena rather than merely as products of anatomical variation.[8,9] For instance, individuals may modulate f_o and formant frequencies along a range of their physiologic capabilities to express a variety of communication styles.[9] Features such as vowel space area (VSA) and /s/ spectral characteristics are not dependent on vocal tract size, but rather articulator positioning,[8,10] and may be easily manipulated to inform style.[11]

There are, however, relatively few studies that have examined the acoustic-phonetic characteristics of non-binary individuals. Moreover, there exist few data on relations among these characteristics and ubiquitous perceptual measures of speaker gender and masculinity/femininity. Understanding which features perceptually distinguish speakers of diverse gender identities would help to clarify the enregisterment of gender diversity in speech communication (i.e., how specific features are recognized by listeners as belonging to specific social groups).[12] For example, the acoustic-phonetic cues that

[a]Author to whom correspondence should be addressed.

listeners reference to perceptually organize speakers may vary depending on the gender they perceive the speaker to be.[13] These data may also provide a reference for speech therapists who work with clients seeking to align their gender expression with their gender identity.

Most acoustic-phonetic measurements of non-binary speakers have been limited to characteristics of /s/ and/or f_o.[7,8] Although /s/ and f_o are stylistically and perceptually important features for communicating speaker gender,[14,15] lower formant frequencies[14,15] and smaller VSA[14,16] tend to be perceived as more male-like and/or masculine. No published studies have examined how formant frequencies or VSA of non-binary speakers may systematically differ from cisgender and transgender men and women. Additionally, some studies have suggested that vocal quality, and specifically perceived breathiness, may be a meaningful perceptual cue to speaker gender.[17,18] Recent work found that, in comparison to cisgender men and women, non-binary speakers produced higher cepstral peak prominence-smoothed (CPPs) values, an acoustic correlate of perceptual breathiness,[19] while maintaining f_o within an intermediate range of cisgender speakers.[5] This finding suggests that non-binary speakers may use vocal quality as a sociolinguistic variable to convey gender identity. However, relations between speakers' CPPs values and their perceived gender were not examined in that study.

Given the dearth of information on acoustic-phonetic and auditory-perceptual characteristics of non-binary speakers, this exploratory study had three goals. The first was to examine how non-binary speakers may leverage a specific composite of acoustic-phonetic features to construct and convey a linguistic style distinct from cisgender and transgender men and women. The specific acoustic-phonetic measurements were chosen based on established relations with auditory-perceptual judgments of speaker gender and included mean f_o; CPPs; first (F1), second (F2), and third (F3) formant frequencies; and VSA.[20] The second goal was to compare listeners' ratings of non-binary speakers along commonly used continuous scales for gender identity and masculinity/femininity to those of cisgender and transgender men and women. These specific scaling methods were selected to allow preliminary comparisons between this study and prior work.[20] The third goal was to assess how well the selected acoustic features could predict listeners' perceptual ratings across the differing speaker groups.

2. Methods

2.1 Speakers

This research took place at the Indiana University Department of Speech, Language, and Hearing Sciences after receiving approval from the Institutional Review Board at Indiana University. In total, 59 monolingual, native American English speakers were recruited between the ages of 18 and 67 [mean = 25.9, standard deviation (SD) = 10.8]. No speakers reported a history of speech, language, or hearing disorder. Speakers self-reported their gender identity using a validated demographic instrument.[21] Transgender speakers included 15 non-binary individuals, 7 transgender men, and 7 transgender women. Eleven of the transgender speakers who identified as either male or non-binary reported taking exogenous testosterone therapy as part of their gender affirming care. Time on testosterone for these speakers at the time of recording ranged from 2.5 months to 4 years, and dosages, when reported, varied. Transgender speakers were age- and dialect-matched with 15 cisgender men and 15 cisgender women. There were no significant age differences among cisgender and transgender speaker groups as determined by one-way analysis of variance (ANOVA), $F(4,\ 54) = 0.957$, $p = 0.44$. Cisgender speakers were primarily taken from the ALLSSTAR corpus ($n = 23$).[22] Seven additional cisgender speakers (two women and five men) were recruited specifically for this study as age and dialect matches to the transgender speakers. All transgender speaker recordings and six cisgender speaker recordings were conducted in the Speech Perception Lab in the Department of Speech, Language, and Hearing Sciences at Indiana University. One dialect match cisgender speaker was recorded in the Department of Communicative Disorders at the University of Alabama.

Speakers completed a demographic and language background questionnaire before completing the speech recordings in a sound-treated booth. Those recorded at Indiana University wore a head-mounted Shure (Niles, IL) Dynamic WH20XLR microphone positioned at approximately 5 cm mouth-to-microphone distance. Speakers were recorded using a Marantz (Kawasaki, Japan) PDM670 digital audio recorder connected to a Mac Mini at a 44.1 kHz sampling rate and 16-bit resolution and downsampled to 22.05 kHz. The dialect match cisgender speaker recorded at the University of Alabama used a Steinberg audio interface and Electro Voice RE-20 dynamic cardioid mic at a 44.1 kHz sampling rate and 16-bit resolution that was downsampled to 22.05 kHz. The remaining 23 cisgender speaker recordings from the ALLSSTAR corpus were recorded at Northwestern University using a Shure SM81 Condenser microphone connected to an Intel Core 2 Duo iMac at a 22.05 kHz sampling rate and 16-bit resolution.

Speakers were recorded reading 60 sentences aloud from the Hearing In Noise Test (HINT).[23] They were instructed to use their habitual speaking rate and most comfortable vocal presentation during recording. Total time to complete the recording protocol was between 1 and 1.5 h with breaks provided as needed. Speakers recorded specifically for this study were paid $10 per hour for their participation.

Acoustic measures were collected for each speaker from the read utterance "They had two empty bottles" and included mean f_o and CPPs across the duration of each speaker's utterance, F1, F2, and F3 of the four corner vowels /i, u, ae, ɑ/ in the words "empty," "two," "had," and "bottles" and VSA. All acoustic analyses were completed using Praat.[24] A description of methods to obtain acoustic-phonetic measurements is provided in the supplementary material.[25]

2.2 Listeners

Listeners were 30 cisgender, monolingual, American English speakers (13 women and 17 men), age 18–40 (mean = 22.9, SD = 5.2), who were recruited from the Indiana University campus and surrounding Bloomington community. All listeners passed a pure-tone hearing screening at 250, 500, 1000, 2000, 4000, and 8000 Hz presented at 20–25 dB. Listeners had completed no more than two courses in communication sciences and disorders and no coursework in speech science or voice disorders.

The rating task took place in the Speech Perception Lab in the Department of Speech, Language, and Hearing Sciences at Indiana University. Listeners were seated at a ViewSonic (Brea, CA) 20 in. VX2033wm LCD monitor in a sound booth and wore Sennheiser HD 280 Pro headphones. Stimuli were intensity scaled to 70 dB, and randomization, presentation, and listener response recording were controlled through PsychoPy 2021.2.3.[26] Listeners first completed speaker sorting tasks that are not discussed further here. Then listeners rated confidence in speaker gender (with anchors at either end for "Definitely male" or "Definitely female") and masculinity/femininity (with anchors at either end for "Very masculine" and "Very feminine"), among other speech dimensions, using nine-point equal appearing interval scales. They were presented with each speaker's utterance one at a time. The order of rating scales and anchor labels was quasi-randomized. Listeners could click onscreen to replay each utterance as often as they liked while completing the ratings. Listeners first completed practice ratings of three speakers, one from each group of cisgender men, cisgender women, and transgender individuals, reading a different HINT sentence from that used in the experiment. Time to complete the protocol was approximately 2.5 h with breaks provided as needed. Listeners were paid $10 per hour for their participation.

3. Results

3.1 Speaker acoustic-phonetic measures

Summaries of acoustic-phonetic measures across speaker groups and statistical model outputs are provided in the supplementary material.[25] The six acoustic measures are plotted in Fig. 1.

To examine group differences for each measure, speaker group (cisgender man, cisgender woman, transgender man, transgender non-binary, and transgender woman) was fit to a linear model using the lm() function of the *stats* package in R.[27] The predictor variable was treatment-coded with cisgender man as the reference level. The overall effect of speaker group was evaluated by comparing the speaker group model to an intercept only model using the anova() function.[28] Inclusion of speaker group significantly improved model fit for each acoustic measure. To assess differences between specific groups, *post hoc* pairwise comparisons were conducted using the *emmeans* package with Holm correction. Only comparisons between non-binary speakers and the other speaker groups are presented here.

Overall, non-binary speakers demonstrated a larger range of values across the six acoustic dimensions than transgender men and transgender women, and their values were generally within an intermediate range of cisgender men and cisgender women. Non-binary speakers' f_o values were higher than those of cisgender men [$\beta = 67.16$, standard error (SE) = 1.79, $t = 37.46$, $p < 0.001$) but lower than those of cisgender women ($\beta = -18.09$, $SE = 1.79$, $t = -10.09$, $p < 0.001$). In comparison to other transgender speakers, non-binary speakers' values were higher than for both transgender men ($\beta = 60.11$, $SE = 2.25$, $t = 26.74$, $p < 0.001$) and transgender women ($\beta = 10.59$, $SE = 2.25$, $t = 4.71$, $p < 0.001$).

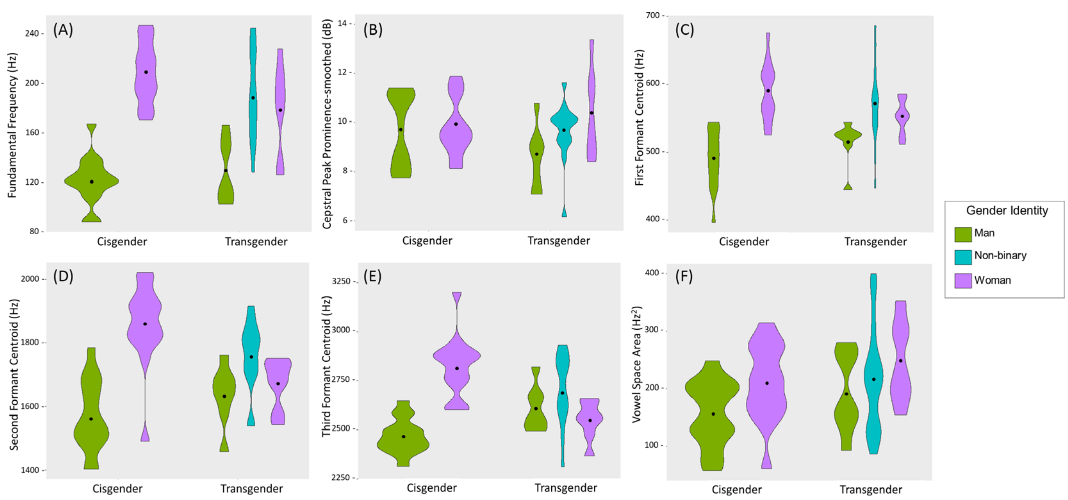

Fig. 1. Violin plots of acoustic measures across speaker groups. Plot widths represent probability density at a given value. Black circles represent mean values for each speaker group. (A) Fundamental frequency (Hz); (B) CPPs (dB); (C) first formant centroid (Hz); (D) second formant centroid (Hz); (E) third formant centroid (Hz); (F) VSA (Hz^2).

This pattern was also seen for the F1 centroid ($\beta = 80$, $SE = 2.88$, $t = 27.75$, $p < 0.001$; $\beta = -18.7$, $SE = 2.88$, $t = -6.47$, $p < 0.001$; $\beta = 55.1$, $SE = 3.62$, $t = 15.24$, $p < 0.001$; and $\beta = 18$, $SE = 3.62$, $t = 4.98$, $p < 0.001$, for cisgender men, cisgender women, transgender men, and transgender women, respectively), F2 centroid ($\beta = 188.9$, $SE = 7.11$, $t = 26.55$, $p < 0.001$; $\beta = -99.3$, $SE = 7.11$, $t = -13.96$, $p < 0.001$; $\beta = 122.6$, $SE = 8.92$, $t = 13.75$, $p < 0.001$; and $\beta = 88.4$, $SE = 8.92$, $t = 9.91$, $p < 0.001$, for cisgender men, cisgender women, transgender men, and transgender women, respectively), and F3 centroid ($\beta = 238.3$, $SE = 8.85$, $t = 26.92$, $p < 0.001$; $\beta = -115.7$, $SE = 8.85$, $t = -13.07$, $p < 0.001$; $\beta = 96.8$, $SE = 11.1$, $t = 8.72$, $p < 0.001$; and $\beta = 160$, $SE = 11.1$, $t = 14.42$, $p < 0.001$, for cisgender men, cisgender women, transgender men, and transgender women, respectively).

No differences in CPPs values were found between non-binary and cisgender men and women speakers. However, non-binary speakers' CPPs values were higher than those of transgender men ($\beta = 1.07$, $SE = 0.11$, $t = 9.81$, $p < 0.001$) but lower than those of transgender women ($\beta = -0.72$, $SE = 0.11$, $t = -6.64$, $p < 0.001$).

Non-binary speakers' VSA was larger than that of cisgender men ($\beta = 59.65$, $SE = 4.67$, $t = 12.78$, $p < 0.001$) but not significantly different from that of cisgender women. Non-binary speakers' VSA was larger than that of transgender men ($\beta = 22.87$, $SE = 5.85$, $t = 3.91$, $p < 0.001$) but smaller than that of transgender women ($\beta = -30.81$, $SE = 5.85$, $t = -5.27$, $p < 0.001$).

3.2 Listener auditory-perceptual measures

Listeners' ratings of confidence in speaker gender and masculinity/femininity for each speaker group are plotted in Fig. 2.

To examine differences in listeners' ratings, speaker group (cisgender man, cisgender woman, transgender man, transgender non-binary, and transgender woman) was fit to a mixed effects model as a fixed effect using the lmer() function of the *lme4* package. The predictor variable was treatment-coded with cisgender man as the reference level. The random effects structure (for this and all following models) included random intercepts by listener and speaker. The overall effect of speaker group was evaluated by comparing the speaker group model to an intercept only model using the anova() function.[28] Inclusion of speaker group significantly improved model fit for both perceptual measures. To assess differences between non-binary speakers and other groups, *post hoc* pairwise comparisons were obtained using the *emmeans* package with Holm correction.

For confidence in gender identification, non-binary speakers were rated as more likely female than were cisgender men ($\beta = -4.48$, $SE = 0.49$, $t = -9.09$, $p < 0.001$), transgender men ($\beta = -3.82$, $SE = 0.62$, $t = -6.18$, $p < 0.001$), and transgender women ($\beta = 2.9$, $SE = 0.62$, $t = 4.69$, $p < 0.001$) but less likely female than were cisgender women ($\beta = -1.85$, $SE = 0.49$, $t = -3.76$, $p < 0.01$). For masculinity/femininity, non-binary speakers were rated as more feminine than cisgender men ($\beta = -3.59$, $SE = 0.37$, $t = -9.81$, $p < 0.001$), transgender men ($\beta = -3.02$, $SE = 0.46$, $t = -6.6$, $p < 0.001$), and transgender women ($\beta = 1.53$, $SE = 0.46$, $t = 3.33$, $p < 0.01$) but as less feminine than cisgender women ($\beta = -1.34$, $SE = 0.37$, $t = -3.66$, $p < 0.01$).

To assess which acoustic features were predictive of confidence in speaker gender and masculinity/femininity for all speakers combined, each auditory-perceptual measure was fit to a mixed effects model with the six acoustic measures as fixed effects. Multicollinearity among acoustic measures was assessed by calculating variance inflation factors using the check_collinearity() function of the *productivity* package. F2 and F3 centroids demonstrated variance inflation factor values > 4 for both response variables and were removed.[27] The final structure for both models was $x \sim f_o + \text{CPPs} + \text{F1 centroid} + \text{VSA} + (1|\text{listener}) + (1|\text{speaker})$. The significance of each acoustic measure was determined through likelihood ratio tests of models with each acoustic measure against models without each acoustic measure using the mixed() function of the *afex* package.[28] For both confidence in speaker gender and masculinity/femininity, only f_o [χ^2 (1) = 37.24, $p < 0.001$, and χ^2 (1) = 52.91, $p < 0.001$, respectively] and F1 centroid [χ^2 (1) = 15.24, $p < 0.001$, and χ^2 (1) = 20.61, $p < 0.001$,

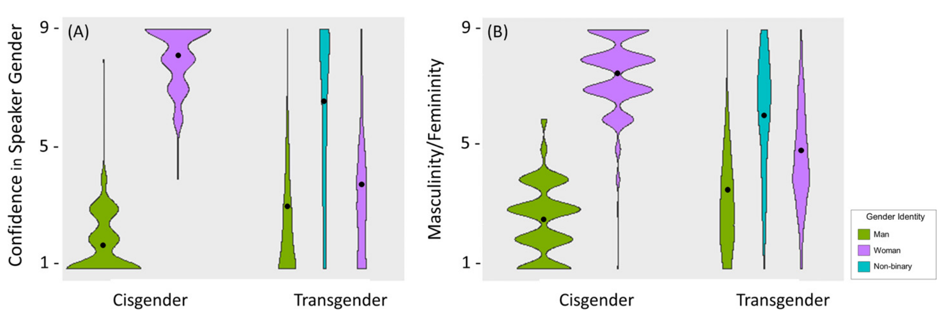

Fig. 2. Violin plots of listeners' auditory-perceptual ratings of (A) confidence in speaker gender and (B) masculinity/femininity. Confidence in speaker gender references listeners' certainty that the speaker is male or female. Masculinity/femininity references listeners' perception that the individual's speech is characteristic of typical men and women. Higher values along the *y* axis indicate (A) greater certainty that the speaker is female and (B) greater perceived femininity. Lower values indicate (A) greater certainty that the speaker is male and (B) greater perceived masculinity. Black circles represent mean values for each speaker group.

respectively] were significant predictors of ratings. The mixed effects models accounted for 86% of the variance for confidence in speaker gender and 78% of the variance for masculinity/femininity.

The above analyses were then conducted on subsets of cisgender vs transgender speakers to examine whether the same acoustic-phonetic features were relevant for perceptual ratings between the two speaker groups. For cisgender speakers, only f_o was a significant predictor of both confidence in speaker gender [χ^2 (1) = 54.63, $p < 0.001$] and masculinity/femininity [χ^2 (1) = 61.4, $p < 0.001$]. The models accounted for 92% and 85%, respectively, of the variance in ratings. For transgender speakers, f_o [χ^2 (1) = 20.8, $p < 0.001$], F2 centroid [χ^2 (1) = 20.59, $p < 0.001$], and CPPs [χ^2 (1) = 4.13, $p < 0.05$] were predictive of confidence in speaker gender. Fundamental frequency [χ^2 (1) = 21.05, $p < 0.001$] and F3 centroid [χ^2 (1) = 5.43, $p < 0.05$] were predictive of masculinity/femininity. The models accounted for 79% and 68%, respectively, of the variance in ratings.

4. Discussion and conclusions

Non-binary speakers' acoustic profile was found to be distinct from those of cisgender and transgender men and women. Although some acoustic measures for non-binary speakers (e.g., f_o and formant centroids) fell within an intermediate range of cisgender men and women, their values were higher than those of transgender men and women. Other measures, such as CPPs and VSA, demonstrated more complex patterns in relation to the other speaker groups. For example, CPPs was found to significantly vary only among non-binary speakers and transgender men (who had the lowest values) and transgender women (who had the highest). These findings relate to those of Brown and Pillow-Loiseau,[5] where non-binary speakers demonstrated significantly higher CPPs values than cisgender men and women. Although the present study did not find significant differences between non-binary and cisgender speakers, the distinction among the transgender groups indicates vocal quality may be a sociolinguistic variable leveraged by non-binary speakers to stylistically distinguish themselves from other social groups. Non-binary speakers' VSA, a measure of articulatory precision,[20] also followed a unique pattern, falling within an intermediate range of transgender men and women and being similar to that of cisgender women. VSA has been found to systematically vary based on speakers' self-reported gender identity[29] and to be perceptually relevant to listeners when evaluating speaker gender.[16] Findings from this prior work in conjunction with those of the present study, thus, support VSA as an articulatory feature that may also be manipulated by speakers to position themselves apart from the gender binary. Collectively, these findings suggest non-binary speakers utilize fine phonetic detail to index gender identity and demonstrate how gender, as an evolving social category, contributes to the acoustic variability in speech. Findings also provide insight into how linguistic choices and social meaning may become intertwined and enregister non-binary identity as a distinct social category for speakers and listeners.[12] Additional research with a larger pool of speakers and more varied speech utterances would better clarify how non-binary speakers may utilize a specific composite of acoustic-phonetic features to index their gender identity. Future work may also improve on the present study by including more than one token per speaker and using consistent recording conditions across all speakers to minimize recording artifact that may influence results.[30]

Listeners' ratings of both confidence in speaker gender and masculinity/femininity situated non-binary speakers between cisgender women and all other groups. These findings demonstrate how listeners perceptually organize non-binary speakers along these commonly used scales and provide insight into the cognitive architecture on which listeners may condition speech perception and representation.[31] For instance, listeners may interpret the same acoustic signal differently depending on the gender they attribute to the speaker.[32] However, although this study's paradigm allows comparison to prior work, some limitations should be noted. First, the number of times listeners replayed each speaker's utterance was not collected. Future work could record these data as a proxy for or complement to confidence ratings. Second, using bipolar rating scales to characterize non-binary speakers may perpetuate the gender binary that these individuals themselves reject. As the speaker data suggest, there is an ideologically emergent category of "non-binary" that may also be maintained by listeners. Restricting listeners to the scales used here may have prompted them to reference stereotypes of a binary model of gender and/or prevented them from demonstrating a non-binary or "neutral" category that exists in their perceptual representation. Third, masculinity and femininity may constitute discrete perceptual dimensions rather than two end points of a single continuum.[33] These limitations stem, in part, from socially conditioned researcher biases surrounding gender. It is, therefore, critical that researchers be reflexive when developing and implementing research with the gender diverse community and acknowledge that their own experiences, assumptions, and beliefs may influence the research process.[34]

Relations among the acoustic-phonetic features and listeners' perceptual ratings showed different patterns both between cisgender and transgender speakers and between the two rating scales. Only f_o and F1 centroid were predictive of ratings for both confidence in speaker gender and masculinity/femininity when all speakers were pooled. This finding may have arisen because listeners were constrained to using the two bipolar scales and conditioned their response to these dimensions on only f_o and formant information. Other perceptual tasks (e.g., rating independent scales of masculinity and femininity or free classification) may better reveal the interaction among additional acoustic-phonetic features and demonstrate greater nuance in how these features inform listeners' representation of speaker gender.[35] However, differing acoustic-phonetic features were predictive of listeners' ratings when analyzing cisgender and transgender speakers independently. Listeners appeared to rely primarily on f_o when rating cisgender speakers but on f_o, formant frequencies, and vocal quality when rating transgender speakers. This discrepancy between the two groups suggests that listeners may have had difficulty relying solely on f_o and formant information when rating transgender speakers. They may have, therefore, utilized additional acoustic-phonetic cues in making their judgments. This possibility is reflected in the fact that more variance in listeners' ratings was explained by fewer acoustic features for cisgender as compared to transgender speakers. These findings may also be partly due to variability across listeners in their perception of speaker gender. Paradigms that are sensitive to individual differences would better clarify how and why listeners may vary in their perceptual representation of gender.

This study adds to the nascent body of research investigating acoustic-phonetic and auditory-perceptual characteristics of gender non-binary speakers. These data provide insight into the acoustic-phonetic features speakers utilize to index non-binary identity and the perceptual relevance of these features to listeners' judgments of speaker gender. Findings suggest a socially grounded account of acoustic-phonetic variability among diverse gender identities and evolving linguistic landscapes to which listeners must adapt.

Acknowledgments

The author is grateful to the Department of Speech, Language, and Hearing Sciences at Indiana University for funding in support of this research. The author also thanks Tessa Bent for her academic guidance in completing this work and three anonymous reviewers for their constructive feedback on an earlier version of the manuscript.

References and links

[1]American Psychological Association, "Guidelines for psychological practice with transgender and gender nonconforming people," Am. Psychol. **70**(9), 832–864 (2015).
[2]C. N. Fiani and H. J. Han, "Navigating identity: Experiences of binary and non-binary transgender and gender non-conforming (TGNC) adults," Int. J. Transgender. **20**(2–3), 181–194 (2019).
[3]J. K. Mcguire, T. F. Beek, J. M. Catalpa, and T. D. Steensma, "The genderqueer identity (GQI) scale: Measurement and validation of four distinct subscales with trans and LGBQ clinical and community samples in two countries," Int. J. Transgender. **20**(2–3), 289–304 (2019).
[4]N. J. Bradford and J. M. Catalpa, "Social and psychological heterogeneity among binary transgender, non-binary transgender and cisgender individuals," Psychol. Sex **10**(1), 69–82 (2019).
[5]L. Brown and C. Pillot-Loiseau, "Bright voice quality and fundamental frequency variation in non-binary speakers," J. Voice (published online 2022).
[6]A. Corwin, "Language and gender variance: Constructing gender beyond the male/female binary," http://mail.ejhs.org/Volume12/Gender.htm (Last viewed October 24, 2022).
[7]L. Zimman, "Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/," Lang. Soc. **46**(3), 339–370 (2017).
[8]L. Zimman, "Variability in /s/ among transgender speakers: Evidence for a socially grounded account of gender and sibilants," Linguistics **55**(5), 993–1019 (2017).
[9]P. Eckert, "Style and social meaning," in *Style and Sociolinguistic Variation*, edited by P. Eckert and J. R. Rickford (Cambridge University, Cambridge, UK, 2001), pp.119–126.
[10]P. Flipsen and S. Lee, "Reference data for the American English acoustic vowel space," Clin. Linguist. Phon. **26**(11–12), 926–933 (2012).
[11]B. Merritt and T. Bent, "Revisiting the acoustics of speaker gender perception: A gender expansive perspective," J. Acoust. Soc. Am. **151**(1), 484–499 (2022).
[12]A. Agha, *Language and Social Relations* (Cambridge University, Cambridge, UK, 2007).
[13]T. Murry and S. Singh, "Multidimensional analysis of male and female voices," J. Acoust. Soc. Am. **68**(5), 1294–1300 (1980).
[14]B. Munson, "The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation," Lang. Speech **50**(1), 125–142 (2007).
[15]Y. Leung, J. Oates, S. P. Chan, and V. Papp, "Associations between speaking fundamental frequency, vowel formant frequencies, and listener perceptions of speaker gender and vocal femininity-masculinity," J. Speech Lang. Hear. Res. **64**, 2600–2622 (2021).
[16]J. A. Booz, "Perceived gender in clear and conversational speech," Master's thesis, University of Utah, Salt Lake City, UT, 2016.
[17]J. Bishop and P. Keating, "Perception of pitch location within a speaker's range: Fundamental frequency, voice quality and speaker sex," J. Acoust. Soc. Am. **132**(2), 1100–1112 (2012).
[18]V. G. Skuk and S. R. Schweinberger, "Influences of fundamental frequency, formant frequencies, aperiodicity, and spectrum level on the perception of voice gender," J. Speech Lang. Hear. Res. **57**(1), 285–296 (2014).
[19]J. Hillenbrand and R. A. Houde, "Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech," J. Speech Lang. Hear. Res. **39**(2), 311–321 (1996).
[20]Y. Leung, J. Oates, and S. P. Chan, "Voice, articulation and prosody contribute to listener perceptions of speaker gender: A systematic review and meta-analysis," J. Speech Lang. Hear. Res. **61**, 266–297 (2018).
[21]M. L. Haupert, A. R. D. Pope, J. R. Garcia, and E. R. Smith, "An inclusive gender identity measure," in *Handbook of Sexuality-Related Measures*, 4th ed. (Routledge, Abingdon, UK, 2019), pp. 353–355.
[22]Bradlow, A. R., "ALLSSTAR: Archive of L1 and L2 scripted and spontaneous transcripts and recordings," https://speechbox.linguistics.northwestern.edu/#!/?goto=allsstar (Last viewed August 14, 2022).
[23]M. Nilsson, S. D. Soli, and J. A. Sullivan, "Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am. **95**(2), 1085–1099 (1994).

[24]P. Boersma and D. Weenink, "Praat: Doing phonetics by computer (version 6.1.38) [computer program]," http://www.praat.org/.
[25]See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0017642 for a detailed description of methods to obtain each acoustic measure and a summary of acoustic-phonetic measures across speaker groups and statistical model outputs.
[26]J. Peirce, J. R. Gray, S. Simpson, M. MacAskill, R. Höchenberger, H. Sogo, E. Kastman, and J. K. Lindeløv, "PsychoPy2: Experiments in behavior made easy," Behav. Res. **51**, 195–203 (2019).
[27]R Core Team. "R: A language and environment for statistical computing," https://www.r-project.org (Last viewed October 1, 2022).
[28]B. Winter, *Statistics for Linguists* (Taylor & Francis, London, 2020).
[29]M. Weirich and A. P. Simpson, "Gender identity is indexed and perceived in speech," PLoS One **13**(12), e0209226 (2018).
[30]J. G. Švec and S. Granqvist, "Guidelines for selecting microphones for human voice production research," Am. J. Speech Lang. Pathol. **19**(4), 356–368 (2010).
[31]L. C. Nygaard and C. Y. Tzeng, "Perceptual integration of linguistic and nonlinguistic properties of speech," in *The Handbook of Speech Perception*, 2nd ed., edited by J. Pardo, L. Nygaard, R. Remez, and D. Pisoni (Wiley-Blackwell, New York, 2021), pp. 390–413.
[32]K. Johnson, E. A. Strand, and M. D'Imperio, "Auditory-visual integration of talker gender in vowel perception," J. Phon. **27**(4), 359–384 (1999).
[33]A. Constantinople, "Masculinity-femininity: An exception to a famous dictum?," Psychol. Bull. **80**(5), 389–407 (1973).
[34]D. Watt, "On becoming a qualitative researcher: The value of reflexivity," Qual. Rep. **12**, 82–101 (2007).
[35]B. Merritt, "Perceptual representation of speaker gender," Ph.D. dissertation, Indiana University Bloomington, Bloomington, IN, 2022.