ACOUSTIC CORRELATES OF SELF-ASCRIBED MASCULINITY/FEMININITY IN AN AMERICAN ENGLISH SAMPLE

Chen Zhou[1], Jason Bishop[1,2]

[1]The Graduate Center, CUNY; [2] College of Staten Island, CUNY
czhou1@gradcenter.cuny.edu, jason.bishop@gc.cuny.edu

ABSTRACT

Recently it has been argued that some aspects of the signal usually associated with a speaker's biological sex may also vary systematically in relation to measures of the speaker's masculinity/femininity, suggesting gender identity, gradiently defined, may be a source of inter-speaker variation in speech. The present study attempted to extend the findings from a recent production study reported in [1] (which tested a German sample) to English. Read speech from a sample of 74 American English speakers (38 women, 36 men) was collected and analyzed with respect to two variables known to vary, in part, as a function of speaker sex and/or gender: vowel space area and the spectral center of gravity for sibilant /s/. These acoustic phonetic variables were then considered in relation to gradient measures of the speakers' self-ascribed femininity. Results were partially and only weakly consistent with the previously-reported German findings.

Keywords: masculinity/femininity, socio-phonetics, gender identity, individual variability in speech.

1. INTRODUCTION

A number of acoustic-phonetic factors are known to distinguish, to varying degrees, the speech of men and women. Some of the most well-studied aspects of the signal believed to distinguish the sexes, biologically defined, include measures of spoken fundamental frequency (f0) – especially mean f0 ([2], [3]) and f0 range ([4]) – as well as differences related to voice quality ([5]), acoustic vowel space size ([6], [7]), and the spectral characteristics of sibilant fricatives ([8], [9]). For example, in comparison to men's speech, women have generally been shown to produce higher mean f0, a wider f0 range, a larger acoustic vowel space area and higher spectral mean (i.e., center of gravity), for sibilant /s/. Importantly, however, such phonetic differences have been described as reflecting not only biophysical differences between the male and female sexes, but also learned differences that distinguish men and woman as gender groups, where 'gender' is understood as a cultural/socio-cognitive construct (e.g., [10], [11]). In other words, one's speech patterns encode information about both sex (biological dimorphisms distinguishing males and females) and gender (learned/abstract social divisions between men and woman).

However, both in the language sciences and in many societies more generally, there is growing recognition that understanding the role that sex and gender play in speech will likely require recognizing more than just these binary, categorical distinctions. On the one hand, there is well documented within-sex variation tied systematically to distinctions in sexual orientation, such as straight versus non-straight speakers ([12], [13]). However, it is also known that gender characterizations not directly related to sexual orientation, such as a coarse distinction between cisgender and gender expansive speakers, have correlates in the speech signal as well ([14], [15]). Most pertinent to our purposes here, there is reason to believe that gender – again a socio-cognitive construct – can be understood as not only non-binary, but perhaps not even categorical. If this is true, then we would expect to see gradient measures of gender or 'gender identity' predict properties of speech outputs in correspondingly gradient ways.

In fact, there is some evidence for this kind of intra-group variability as well. For example, two studies testing German speakers ([16], [17]) found that some acoustic cues associated with sex-specific differences varied as a function of speakers' scores on continuous self-report scales of masculine/feminine social/personality characteristics, which have been claimed to capture fundamental aspects of a person's sense of gender identity ([18]). Moreover, this was true within groups of lesbian women (in terms of median f0) as well as within groups of straight and gay men (in terms of mean second formants for /aː/, /iː/ and /uː/). Such findings suggest the possibility that previous work may have overestimated the role of sex in determining some phonetic properties of speech. Instead, a significant amount of variation we see might actually be due to individual differences along some socio-cognitive dimension(s) related to gender identity.

Few studies have explored the effect of such individual variability in gender identity on sex- and gender-specific speech patterns. One exception was [1]'s investigation of German speakers. In that study, gender identity was operationalized in terms of self-reported ratings on a scale of masculinity/femininity and the authors examined the relationship between these ratings and several acoustic parameters known to encode sex and gender. Two questionnaires were used to elicit speakers' femininity ratings: the F+ scale of the German version of the Extended Personal Attributes Questionnaire (GEPAQ-F+) ([19]) and the Traditional Masculinity-Femininity scale (TMF) ([20]). The authors found that both vowel space area (VSA) and mean f0 varied systematically in relation to self-ascribed femininity ratings, though only for male speakers. More specifically, the more feminine/less masculine that men rated themselves to be, the larger their VSA and the higher their f0 were (i.e., in the direction of more female-like values). The other exception was [21]'s study on the moderating effect of social meaning on pitch range differences of female and male Japanese-British English sequential bilinguals. Both female and male speakers' English pitch level (measured with mean f0) were found to vary as a function of individual gender identity, also operationalized as scores on continuous scales of masculinity and femininity (in their case, using the short version of the Bem Sex Role Inventory (BSRI-short). The findings from both studies invite further examination of subtle-but-systematic acoustic cues to constructs such as gender identity, which can arguably be usefully operationalized in terms of self-ascribed masculinity/femininity.

To this end, the present study attempted to partially replicate the basic findings reported by [1], who found gradient measures of femininity to be systematically related to subtle differences in speech usually associated with sex. Our goal, however, was to extend the discussion to another language, namely American English, and to use a larger sample size, as [1]'s data set was rather modest for evaluating interindividual variation tied to what is, again, best regarded as a socio-cognitive variable. Moreover, we also sought to control for prosodic position and segmental context in the speech samples, which led us to utilize read speech rather than spontaneous or unscripted speech (as in [1]'s study). In line with the investigation in [1], the current study treated gender identity as a continuous variable, based on speakers' self-reported scores on established scales of femininity. With these basic considerations in mind, the present study asked two research questions:

1.) Do speakers with higher self-ratings of femininity produce more female-like values for phonetic variables like VSA or spectral center of gravity (CoG) for /s/? These two variables have been shown previously to vary significantly based on speaker sex (e.g., [22] for VSA; [8], [9] for CoG).

2.) If so, are these gender-dependent patterns to any extent also sex-specific?

Based primarily on the findings in [1] and [21], we made the following predictions. First, we predicted that the more feminine a speaker rates themselves, the more female-like their productions of the acoustic variables will be; in the present case, that would mean larger VSA and higher spectral CoG for /s/. Second, we predicted that the patterns may be specific to a particular sex group. That is, the systematic relationship between the acoustic variables examined and the self-reported femininity scores of the speakers, if any such relationships exist, might be limited to just one sex group, as this was what was found previously by [1] for VSA and mean f0, which were affected by femininity scores for male speakers only. (See also [21], who found gender ratings to have the opposite relationship to mean f0 for male and female speakers).

2. EXPERIMENT

2.1. Methods

2.1.1. Participants

A total of 74 monolingual American English speakers (38 female speakers, 36 male speakers, mean age 21.8 years, SD 3.7) participated in the current study, a previous analysis of which was reported in [23]. In order to limit dialectal variation in the speech sample, all participants recruited in the present study originated and lived all or most of their lives in New York City, and more specifically, Staten Island. None of the participants reported any history of a speech or hearing disorder. When asked to choose options about their self-identified (categorical and binary) gender, only one participant (a biological female who identified as a man in gender) chose an option that was different from their biological sex.

2.1.2 Stimuli

Two sets of English CVC wordforms, consisting of sixteen target words, were used in the study. One set contained eight words eliciting the four vowels /i/, /ɛ/, /u/, and /ɑ/ in two consonantal contexts: h_d and b/p_t. (While the /æ/ vowel is in fact more peripheral than /ɛ/ in American English, it was not used to construct vowel spaces here due to dialectal variation associated with 'a-tensing' in NYC English; [30]). The other set of CVCs consisted of four words intended to elicit /s/ in onset and coda positions. Target words were to be presented in carrier sentences in which the target word was contrasted with another CVC (filler) word in order to induce contrastive/emphatic prosodic prominence. The phrase positions of target words and their relative order with the filler words were also varied to balance the prosodic effects associated with sentence position, yielding four positions for the target words in four carrier sentences:

(1, 2) *Target/filler* is the first word, and *target/filler* is the second word. (Phrase-initial)

(3, 4) The first word is *target/filler*, and the second word is *target/filler*. (Phrase-final)

Stimuli were pseudo-randomized and presented in two blocks whose order was counterbalanced across participants; one block contained phrase-initial targets and fillers and the other block contained phrase-final targets and fillers. One production of each item in each sentence was collected, resulting in 96 samples per speaker and thus 7,104 tokens for acoustic analysis in total.

2.1.3. Measurement of gender identity

Similar to the approach adopted in [1], gender identity was operationalized as scores on two continuous measures of self-reported femininity, also used in [1]: the Femininity scale of the Extended Personal Attributes Questionnaire (EPAQ-F+) ([25]) and the Traditional Masculinity-Femininity Scale (TMF) ([20]), both of which are intended to assess feminine traits and behaviours in English-speaking adults. For each of the questionnaires, the higher the score, the higher the self-rated femininity.

2.1.4. Procedure

All the recordings were made in a sound-attenuated booth using a head-mounted Shure SM-10A microphone and a 44.1kHz sampling rate. Participants were seated in the sound booth in front of a computer screen, on which the target and filler words were presented in their carrier sentences, one carrier sentence at a time on the screen. To create an interactive speaking context for the task, the participants were told to read the carrier sentences for the benefit of a confederate, who sat with the participant in the sound booth (but behind a curtain) so as to write down the target and filler words in each sentence produced by the participant. The confederate would say “got it” to give verbal confirmation after each sentence was produced before moving on to the next. Recordings were saved as *wav* files for later analysis in Praat (version 6.1.15, [26]).

2.2. Acoustic and statistical analyses

As described above, the two acoustic measures in the present study included: (1) the spectral CoG of sibilant /s/, and (2) VSA, the geometric area of the quadrilateral enclosed by the points in an F1 × F2 space whose coordinates were defined by the values (transformed to a Bark scale) of the first two formants for each of the four vowels (/i/, /ɛ/, /u/, and /ɑ/). For each of the acoustic measures, Linear Mixed-Effects (LME) models were used to identify any effects related to self-rated femininity scores in predicting CoG for /s/ and VSA in speakers’ productions.

3. RESULTS

3.1. Self-rated femininity ratings

We first consider the distributions of scores on the two gender-related measures, their relation to each other and how well they each distinguish the two sex groups. To that end, the scatterplot in Fig1 shows the relationship between TMF and EPAQ-F+ for male and female speakers. Simple Pearson correlations showed a marginally significant positive relationship between TMF and EPAQ-F+, but for female speakers only. The box and whisker plots in Fig1 illustrate the amount of overlap in EPAQ-F+ and TMF scores exhibited by the two sex groups; Welch Two sample t-tests for the inter-sex differences on EPAQ-F+ and TMF scores showed a significant sex group difference for TMF scores (t = 15.716, df = 65.751, p < 0.0001), but not for EPAQ-F+ scores (t = 1.6404, df = 68.73, p = 0.1055). Therefore, only the TMF gender scale clearly differentiated the two sex groups in this sample.

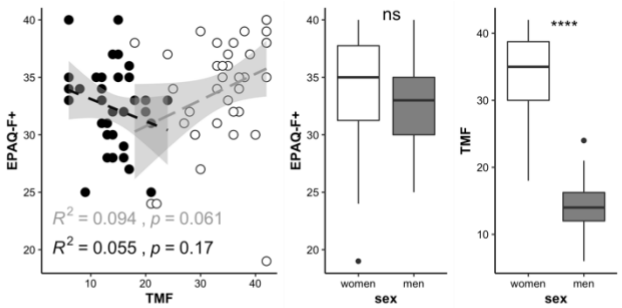

Figure 1. Plots for EPAQ-F+ and TMF scores (male speakers: black; female speakers: white). Left plot: the relationship between the two measures for the two sexes. Right plots: boxplots show the distribution of EPAQ-F+ and TMF scores for each sex.

3.2. Effects of femininity ratings on CoG

Fig2 (left two plots) shows the effect of each measure of femininity on CoG values for /s/. For TMF, though the relationship was nominally in the predicted

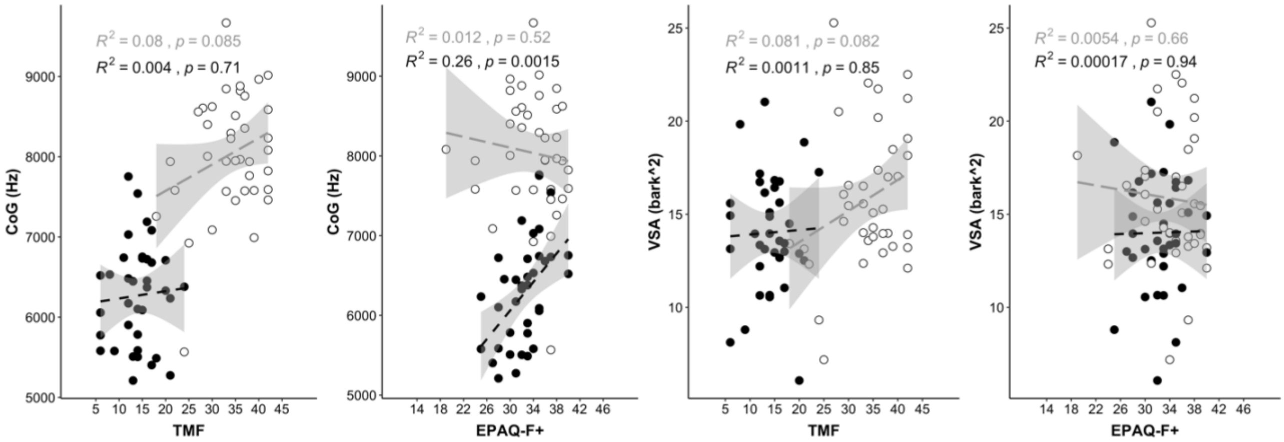

Figure 2. Scatter plots that show the relationships between the femininity scores and the acoustic variables. (Male speakers: black; female speakers: white/grey). Left two plots: the femininity scores in relation to CoG; right two plots: the femininity scores in relation to VSA.

direction for both sexes, a simple Pearson correlation suggested this relationship was weak – only marginally significant, and only in the case of female speakers. EPAQ-F+, on the other hand, was found to have a significant correlation, and in the predicted direction, with CoG values – but for male speakers only. LME analysis further confirmed these patterns; models showed a significant interaction between EPAQ-F+ and sex (χ^2 (1) = 12.01, p < 0.001), as well as a significant main effect for TMF (χ^2 (1) = 7.20, p = 0.007). Separate sex-specific models showed that TMF was only marginally significant in predicting female speakers' CoG values (χ^2 (1) = 3.16, p = 0.075) and that EPAQ-F+ significantly predicts male speakers' CoG values (χ^2 (1) = 12.49, p < 0.001).

3.3. Effects of femininity ratings on VSA

Fig2 (right two plots) illustrate the relationships between the measures of femininity and VSA. There was a weak positive correlation between TMF and VSA, though for female speakers only. This relationship was shown to hold in a more rigorous LME modelling; TMF was found to be a marginally significant predictor of VSA in a sex-specific model of female speakers (intercept = 0.17, SE = 0.09, t = 1.805, p = 0.08), but not for male speakers' VSA (intercept = 0.03, SE = 0.12, t = 0.25, p = 0.80).

4. DISCUSSION/CONCLUSION

The present study examined the relationship between gradiently defined gender identity, operationalized as scores on two continuous scales of self-reported femininity, and two acoustic phonetic variables usually associated with differences in speaker sex. Statistical analyses indicated that, at least for male speakers, higher scores on EPAQ-F+ (indicating higher levels of femininity) were significantly associated with productions of /s/ with higher, more female-like CoG values. This is, in fact, the predicted relationship if we assume that (a) gender can be defined gradiently between two extremes, and (b) speakers encode this subtle information about gender in their speech. Interestingly – and as in [1] – this relationship only held for male speakers, although in the present case, for a different acoustic variable, and only when EPAQ-F+ was the measure of femininity.

However, the findings related to EPAQ-F+ in the present study suggest that that measure may not be tapping into the same gender-related construct as TMF, at least not for our American English sample. That is, unlike GEPAQ-F+ scores in the German sample in [1], EPAQ-F+ scores in the present study did not clearly differentiate male and female speakers, and were also not well correlated with TMF within-sex (see Fig1, above). One possibility, though speculative, is suggested by [20], who note that some well-established femininity/masculinity scales tap into socio-cognitive constructs more closely related to instrumentality or expressivity of gender traits rather than gender identity per se. Further research is therefore likely needed to evaluate this measure. For the time being, we think that the results related to TMF in the present study are likely the most reliable, and therefore that the role of gradiently-defined gender identity was detectable in the speech of this American English sample (for both acoustic variables) – but only marginally so, and only for female speakers.

7. REFERENCES

[1] M. Weirich and A. P. Simpson, "Gender identity is indexed and perceived in speech," *PLOS ONE*, vol. 13, no. 12, p. e0209226, Dec. 2018, doi: 10.1371/journal.pone.0209226.

[2] K. N. Stevens, *Acoustic phonetics*, vol. 30. MIT press, 2000.

[3] I. R. Titze, "Physiologic and acoustic differences between male and female voices," *The Journal of the Acoustical Society of America*, vol. 85, no. 4, pp. 1699–1707, Apr. 1989, doi: 10.1121/1.397959.

[4] J. Haan and V. J. van Heuven, "Male vs. female pitch range in Dutch questions," presented at the Proceedings of the 14th International Congress of Phonetic Sciences, NN, 1999, pp. 1581–1584.

[5] C. G. Henton and R. A. W. Bladon, "Breathiness in normal female speech: Inefficiency versus desirability," *Language & Communication*, vol. 5, no. 3, pp. 221–227, Jan. 1985, doi: 10.1016/0271-5309(85)90012-6.

[6] J. Hillenbrand, L. A. Getty, M. J. Clark, and K. Wheeler, "Acoustic characteristics of American English vowels," *The Journal of the Acoustical Society of America*, vol. 97, no. 5, pp. 3099–3111, May 1995, doi: 10.1121/1.411872.

[7] M. Weirich and A. P. Simpson, "Differences in acoustic vowel space and the perception of speech tempo," *Journal of Phonetics*, vol. 43, pp. 1–10, 2014.

[8] K. Heffernan, "Evidence from HNR that/s/is a social marker of gender," *Toronto Working Papers in Linguistics*, vol. 23, 2004.

[9] A. P. Simpson and M. Weirich, "Gender-specific differences in sibilant contrast realizations in English and German.," in *ICPhS*, 2015.

[10] B. Munson and M. Babel, "The phonetics of sex and gender," in *The Routledge Handbook of Phonetics*, Taylor and Francis, 2019, pp. 499–525.

[11] A. P. Simpson and M. Weirich, "Phonetic Correlates of Sex, Gender and Sexual Orientation," *Oxford Research Encyclopedia of Linguistics*, Oct. 27, 2020. http://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-749 (accessed May 12, 2021).

[12] B. Munson, E. C. McDonald, N. L. DeBoe, and A. R. White, "The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech," *Journal of Phonetics*, vol. 34, no. 2, pp. 202–240, Apr. 2006, doi: 10.1016/j.wocn.2005.05.003.

[13] J. B. Pierrehumbert, T. Bent, B. Munson, A. R. Bradlow, and J. M. Bailey, "The influence of sexual orientation on vowel production (L)," *The Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 1905–1908, Oct. 2004, doi: 10.1121/1.1788729.

[14] L. Zimman, "Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men," *Journal of Language and Sexuality*, vol. 2, no. 1, pp. 1–39, 2013.

[15] L. Zimman, "Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/," *Language in Society*, vol. 46, no. 3, pp. 339–370, Jun. 2017, doi: 10.1017/S0047404517000070.

[16] S. Kachel, A. P. Simpson, and M. C. Steffens, "Acoustic correlates of sexual orientation and gender-role self-concept in women's speech," *The Journal of the Acoustical Society of America*, vol. 141, no. 6, pp. 4793–4809, Jun. 2017, doi: 10.1121/1.4988684.

[17] S. Kachel, A. P. Simpson, and M. C. Steffens, "'Do I Sound Straight?': Acoustic Correlates of Actual and Perceived Sexual Orientation and Masculinity/Femininity in Men's Speech," *Journal of Speech, Language and Hearing Research (Online)*, vol. 61, no. 7, pp. 1560–1578, Jul. 2018, doi: http://dx.doi.org.ezproxy.gc.cuny.edu/10.1044/2018_JSLHR-S-17-0125.

[18] J. T. Spence and C. E. Buckner, "Instrumental and Expressive Traits, Trait Stereotypes, and Sexist Attitudes: What Do They Signify?," *Psychology of Women Quarterly*, vol. 24, no. 1, pp. 44–53, Mar. 2000, doi: 10.1111/j.1471-6402.2000.tb01021.x.

[19] T. E. Runge, D. Frey, P. Gollwitzer, R. L. Helmreich, and J. T. Spence, "Masculine (Instrumental) and Feminine (Expressive) Traits: A Comparison between Students in the United States and West Germany," *Journal of Cross-Cultural Psychology*, vol. 12, no. 2, pp. 142–162, 1981, doi: 10.1177/0022022181122002.

[20] S. Kachel, M. C. Steffens, and C. Niedlich, "Traditional Masculinity and Femininity: Validation of a New Scale Assessing Gender Roles," *Front. Psychol.*, vol. 7, 2016, doi: 10.3389/fpsyg.2016.00956.

[21] E. Passoni, E. de Leeuw, and E. Levon, "Bilinguals Produce Pitch Range Differently in Their Two Languages to Convey Social Meaning," *Language and speech*, vol. 65, no. 4, pp. 1071–1095, 2022.

[22] C. Henton, "Cross-language variation in the vowels of female and male speakers," in *Proc. XIIIth ICPhS*, 1995, pp. 420–423.

[23] J. Bishop *et al.*, "Brief Report: Autistic Traits Predict Spectral Correlates of Vowel Intelligibility for Female Speakers," *J Autism Dev Disord*, May 2021, doi: 10.1007/s10803-021-05087-5.

[24] K. Becker and A. W. Wong, "The Short-a System of New York City English: An Update," *University of Pennsylvania Working Papers in Linguistics*, vol. 15, no. 2, Mar. 2010, [Online]. Available: https://repository.upenn.edu/pwpl/vol15/iss2/3

[25] J. T. Spence and R. L. Helmreich, *Masculinity and femininity: Their psychological dimensions, correlates, and antecedents*. University of Texas Press, 1979.

[26] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [Computer program]. Version 6.1. 15," 2020.