Bark and Hz scaled F2 Locus equations: Sex differences and individual differences

Frank Herrmann[a], Stuart P. Cunningham[b] & Sandra P. Whiteside[c]

[a] Department of English, University of Chester, UK;
[b,c] Department of Human Communication Sciences, University of Sheffield, UK
f.herrmann@chester.ac.uk; s.cunningham@sheffield.ac.uk; s.whiteside@sheffield.ac.uk

ABSTRACT

This study investigated speaker sex differences in F2 Locus equations (F2 LEs) based on linearly (Hz) and tonotopically (Bark) scaled formant measurements. F2 data based on English monosyllabic words produced by thirteen women and eleven men were tonotopically scaled [20, 21] and F2 LEs were derived for both the linear and tonotopically scaled formant values. Although the overall sex difference in the F2 LE slope values for women and men was significant for both sets of F2 measures, the magnitude of this difference decreased for the Bark (.047) compared to the Hz (.063) scale.

The individual data revealed a significant correlation between the slope values of the Hz and Bark scale [r = .974; p<.0001] suggesting a lawful relationship between the two metrics. Further probing revealed that the F2 LE data from women were affected more by the Bark conversion than the data from men.

Keywords: sex differences, coarticulation, F2 Locus equations, tonotopic scale.

1. INTRODUCTION

The perceptual relevance of second formant frequency locus equations (F2 LEs) has been investigated and discussed in the literature [4, 17]. However, to our knowledge, there are no reports on F2 LEs based on tonotopically scaled formant values. There is some evidence for sex differences in F2 LEs, with women displaying shallower slope functions than their male peers, which may be associated with lower levels of coarticulation [8, 14]. This study investigated speaker sex differences in F2 LEs as a function of linearly (Hz) versus tonotopically (Bark) scaled formant data.

1.1 F2 Locus equations

F2 LEs are linear regression functions based on a large number of second formant measurements at vowel onset and midpoint across a number of vowel contexts [11] and can be formulated as shown in Eq. (1); with k indicating the slope of the regression function and c the y-intercept.

$$(1) \qquad F2_{onset} = k \times F2_{midpoint} + c$$

Sussman and colleagues discussed the potential use of F2 LEs for stop-place coding in both speech production [18, 19] and speech perception [4, 17]. The relevance of F2 LE slopes for stop-place coding has been debated, as the exposure to single stimuli, i.e. one F2onset-F2midpoint pair, cannot yield slope values [1].

Comparisons of articulatory and acoustic data on the degree of coarticulation have shown that the relationship between the slopes of F2 LEs and articulatory data is not linear, but lawful [12, 13].

1.2 Sex differences

Sex differences in F2 LEs have been mainly associated with differences in y-intercepts, as data from women trended for higher y-intercept values [18, 19]. However, some evidence suggests that speech data from women yields significantly shallower slopes [8, 14].

The smaller degrees of lower level formant variation observed for speech data obtained from men [22] and the overall smaller vowel spaces [6, 10, 20, 22] suggest smaller ranges of F2 values at both vowel onset and midpoint position. This in turn may result in steeper F2 LE slopes. Along with the possible differences in articulatory movement synchronization [15, 16], this may explain the steeper slopes reported for male speech data in previous studies [8, 14]. Data from the latter studies also showed greater F2 ranges for women compared to men. In turn, steeper slopes also go some way to explain the significantly lower y-intercepts for male speech data reported earlier [18, 19]. Overall higher second formant values for front vowels found for female speakers [2] also contribute to the higher y-intercept values. Tonotopic scaling of formant values affects higher frequencies more than lower frequencies [20, 21]. Therefore, Bark scaled F2 LEs should be affected more by the scaling process for speech data from women than men.

This study aimed to compare the magnitude of speaker sex differences in F2 LEs based on linearly (Hz) and tonotopically (Bark) scaled formant measures. Studies on sex differences in other acoustic measures, such as fundamental frequency

[5, 7], have shown that the magnitude of sex differences in F0 ranges may decrease or vanish when using perceptually scaled F0 values. Similarly, it was hypothesized that the differences in the F2 LE slope values between women and men were likely to decrease.

2. METHODS

2.1 Data collection

Twenty-four speakers of British English (13 women, 11 men) repeated eight monosyllabic stimuli after an auditory presentation of the stimulus during 2.5 s inter-stimulus intervals. The stimuli contained predominantly front vowels /bɪz, bɪs, bɪt, bɪb, bɪd, bɛk, bɛst/ and one back vowel /buːst/. Ten repetitions of each item were recorded before the presentation of a new stimulus in the recording series, resulting in eighty sound files per participant (total N = 1,920).

2.2 Data analysis

Incorrect or incomplete productions (e.g. truncated by the inter-stimulus interval) were discarded as well as those speech samples, which did not warrant an acoustic analysis due to voice quality (whisper/ breathy). These amounted to 119 instances, leaving a total of 1,801 sound files for analysis. The second formant was measured at the vowel onset, which was defined as the first glottal pulse after the release of the bilabial plosive, and at the temporal midpoint of the vowel.

Tonotopic scaling of the linear F2 values (Hz) was carried out using the formula (see Eq. 2) described in [1, 2].

(2) $Z = [26.81F/(1960+F)] - 0.53$

Subsequently, sets of F2 LEs were derived (see Eq. 1) based on the linearly and tonotopically scaled F2 formant values.

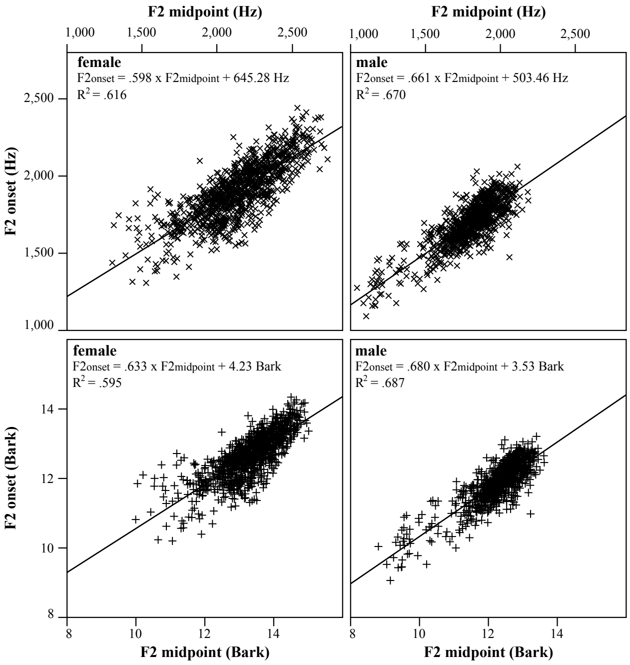
3. RESULTS

3.1 Sex differences

The first part of the analysis focused on overall speaker sex differences. The distribution of the F2 measurements in Hz and Bark at vowel onset and midpoint is shown in the scattergraphs in Fig. 1. The F2 LEs that served for these comparisons are based on the F2 onset and midpoint measurements for all speakers grouped by speaker sex (all men versus all women) and are also displayed in Fig. 1.

A large sample Z-test for parallelism [9] on the Bark scaled formant data revealed that the men exhibited significantly steeper slopes than the women (difference = .047 [Z = 2.031; p<.05]).

Figure 1: Comparison of the F2 Locus equations for the women (left, N = 974) and men (right, N = 974) based on Hz (top) and Bark (bottom) values.



The difference in correlation coefficients underlying the R^2 values was tested using Fisher's z_r transformation test [3] and showed that the men displayed significantly higher correlation coefficients (r = .829) than the women (r = .771; difference = .058 [z_r = 3.423; p<.001]).

These differences in the Bark scaled F2 LEs were comparable to the Hz scaled F2 LEs: significantly steeper slopes for the men (0.661) compared to the women (0.598) [Z = 2.84, p < .005], as well as significantly different correlation coefficients underlying the R^2 values for the male (r = 0.819) and female (r = 0.785) data [z_r = 2.02, p < .05].

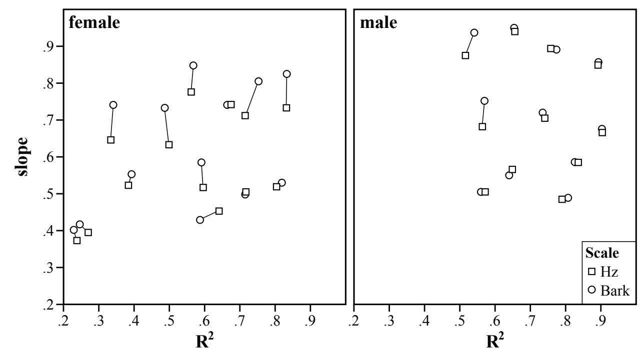
However, closer inspection of Fig. 1 reveals that the differences in slope values between women and men were smaller for the Bark scaled data (.047) compared to Hz scaled data (.063). This suggests that although the speaker sex differences were maintained, the magnitude of this disparity decreased when using the tonotopic (Bark) scale.

3.2 Bark and Hz scaled F2 Locus equations: Individual speaker differences

F2 Locus equations were derived for all twenty-four speakers individually, for both the Hz and Bark scale formant data, resulting in forty-eight F2 LEs. There was a significant correlation [r = .974; p<.0001 (2-tailed)] between the slope values of the Hz and the Bark scale. There was also a significant correlation between the y-intercepts [r = .981; p<.0001 (2-tailed)] as well as the R^2 values [r = .996; p<.0001 (2-tailed)] of the Hz and Bark scaled F2 LE data. This may be indicative of a lawful relationship between the two metrics.

However, closer inspection revealed that not all individual speakers' F2 LE data were affected to the same degree. Fig. 2 demonstrates these individual differences in a *slope x* R^2 space for both women (left) and men (right).

Figure 2: Comparison of Bark (circles) and Hz (squares) based slope and R^2 values of the F2 LEs derived for individual speakers grouped by speaker sex (women: left, men: right).



With regard to the overall distribution of the individual data, the ranges of the R^2 values are greater for the women (Hz: 0.594; Bark: 0.604) compared to the men (Hz: 0.388; Bark: 0.362), which suggests greater variability across the female than male participants. The ranges of the slope values, on the other hand, are similar for both men (Hz: 0.455; Bark: 0.461) and women (Hz: 0.403; Bark: 0.446). However, the slope values of the women appear to be affected most by the scaling process compared to the men, with some of the women's slope values showing marked increases.

3.3 Post-hoc analysis: Slope differences (Bark – Hz)

The slope difference (Bark – Hz) was subjected to a post-hoc analysis to further investigate the individual differences highlighted above. Given the greater peripheral vowel space found for women compared to men [22], the non-uniform sex differences which are more pronounced for front vowels [2, 20], and that higher F2 values are affected to a greater degree by the tonotopic scaling process, the absolute differences between the F2 onset and midpoint values (ΔF2 = |F2midpoint – F2onset|) for the front vowel /ɪ/ in this dataset were further investigated. As the higher F2 values for the female speakers in Fig. 1 cluster closer to the regression line for the Bark F2 LE than the F2 LE based on Hz values, it was hypothesized that the ΔF2 values would positively correlate with the slope differences (Bark – Hz). Participants with greater ΔF2 values were expected to exhibit more pronounced differences between the slope values based on Hz and Bark values, i.e. a greater increase in slope steepness after Bark conversion, compared to participants with comparatively smaller ΔF2 values.

The analysis revealed that there was a significant positive correlation between the slope difference (Bark – Hz) and the mean absolute formant change in F2 (Hz) for the front vowel /ɪ/ [r = .665, p<.01 (2-tailed)]. As was hypothesized, participants with greater ΔF2 values for the front vowel /ɪ/ displayed greater increases in F2 LE slope values after Bark conversion. Further probing revealed that this correlation was only significant for the women [r = .613, p<.05 (2-tailed)], but not the men [r = .520, p>.05 (2-tailed)].

4. DISCUSSION

The comparison of the F2 LEs based on linearly (Hz) and tonotopically (Bark) scaled formant values for all women and men revealed a reduced, but still significant, speaker sex difference in slope values (difference women vs. men: Hz = .063; Bark = .047). This suggests that the steeper F2 LE slope, which may be associated with a greater degrees of coarticulation, previously attributed to men when compared to women [8, 14] is maintained in F2 LEs based on tonotopically scaled formant values, albeit of a reduced magnitude. Previous comparisons of F0 ranges based on Hz and semitone scaled data from women and men similarly revealed that sex differences may decrease on a perceptually motivated scale [5, 7 for German].

The significant correlations between Hz and Bark scaled F2 LE data (slopes, y-intercepts, and R^2 values) derived for individual speakers suggest a close relationship. For most participants an increase in slope values can be observed, which is more pronounced for some speakers compared to others and overall the increase is of a greater magnitude for the women than the men.

Closer inspection of the individual differences in F2 LE slopes based on Hz and Bark scaled formant values (Fig. 2) revealed that the data from women were affected more by the scaling process than the data from men. This may be explained by the Bark conversion, which affects higher frequencies more than lower frequencies [20, 21]. In a similar vein, the phonetic makeup of the current dataset may have contributed to this effect, as it comprised predominantly of high front vowels /ɪ, ɛ/ and, therefore, higher F2 values (as discussed in [8]). These differences in linearly scaled F2 values between women and men may go some way in explaining why women's formant data were affected more by the Bark conversion than men's.

This was further investigated in a post-hoc analysis of the absolute formant change in F2 (ΔF2 in Hz) for the vowel /ɪ/ in section 3.3, as sex differences in F2 values are more pronounced for front than back vowels. Just focusing on all /ɪ/ productions in the current dataset, women exhibited higher ΔF2 values (M = 219.98 Hz, SE = 20.75) than men (M = 121.57, SE = 16.01) and the mean difference (98.41 Hz) was significant [$t(22) = 3.65$, $p<.001$, $r = .61$]. There was a significant correlation between the ΔF2 values for /ɪ/ and the slope differences (Bark – Hz) for the women, but not the men. Tonotopically scaling the F2 onset and midpoint values brings these two values relatively closer together for the women than the men as F2 values are overall higher and are affected more by the scaling process. A further post-hoc comparison of the scaling effect on ΔF2 (100 x (Bark/Hz)) for women (M = 0.316, SE = 0.005) and men (M = 0.361, SE = 0.003) revealed that the mean difference (0.045) was significant [$t(22) = 6.97$, $p<.0001$, $r = .83$], showing the F2 values of the women are affected more by the scaling process. A result of this change in the relative F2 onset – F2 midpoint distance are steeper slopes of the Bark scaled F2 LEs. This may account for the individual differences in the F2 LEs, as speakers with higher ΔF2 values and overall higher F2 values also exhibited the greatest slope differences between the two scales.

5. CONCLUSION

The analysis of this dataset, which comprises predominantly front vowels following word-initial plosives, revealed that sex differences in F2 LEs are reduced when these are calculated based on tonotopically scaled formant values (Bark) instead of linearly scaled formant values (Hz). The correlation between the F2 LE data based on Hz and Bark values for the individual speakers suggests a lawful relationship between the two metrics. The individual differences suggest that the phonetic context, in particular the absolute difference in F2 for the front vowels in this dataset, drives the change in F2 LE slopes. Further research exploring this relationship for other places/manners of articulation of the surrounding consonants as well as a more balanced set of vowel qualities is warranted.

6. REFERENCES

[1] Brancazio, L., Fowler, C. A. 1998. On the relevance of locus equations for production and perception of stop consonants. *Perception & Psychophysics* 60, 24–50.

[2] Fant, G. 1975. Non-uniform vowel normalization. *STL-QPSR* 16, 1–19.

[3] Ferguson, G. A. 1959. *Statistical analysis in psychology and education* (5th ed.) London: McGraw-Hill.

[4] Fruchter, D., Sussman, H. M. 1997. The perceptual relevance of locus equations. *J. Acoust. Soc. Am.* 102, 2997–3008.

[5] Henton, C. G. 1989. Fact and fiction in the description female and male pitch. *Language and Communication* 9, 299–311.

[6] Henton, C. G. 1995. Cross-language variation in the vowels of female and male speakers. *Proc. 13th ICPhS* Stockholm, 420–423.

[7] Herbst, L. 1969. Die Umfänge der physiologischen Hauptsprechtonbereiche von Frauen und Männern. *Zeitschrift für Phonetik und Sprachliche Kommunikation* 22, 426–438.

[8] Herrmann, F., Cunningham, S., Whiteside, S. P. 2014. Speaker sex effects on temporal and spectro-temporal measures of speech. *J. Int. Phon. Assoc.* 44, 59–74.

[9] Kleinbaum, D. G., Kupper, L. L. 1978. *Applied regression analysis and other multivariate methods* Boston: Duxbury Press.

[10] Lee, S., Potamianos, A., Narayanan, S. 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *J. Acoust. Soc. Am.* 105, 1455–1468.

[11] Lindblom, B. 1963. Spectrographic study of vowel reduction. *J. Acoust. Soc. Am.* 35, 1773–1781.

[12] Lindblom, B., Sussman, H. M. 2012. Dissecting coarticulation: How locus equations happen. *J. Phon.* 40, 1–19.

[13] Löfqvist, A. 1999. Interarticulator phasing, locus equations, and degree of coarticulation. *J. Acoust. Soc. Am.* 106, 2022–30.

[14] McLeod, A., Baillargeon, M., Metz, D. E., Schiavetti, N., Whitehead, R. L. 2001. Locus equations as a source of relational invariance for stop place categorization: A direct replication. *Contemp. Issues Commun. Sci. Disord.* 28, 98–103.

[15] Simpson, A. P. 2001. Dynamic consequences of differences in male and female vocal tract dimensions. *J. Acoust. Soc. Am.* 109, 2153–2164.

[16] Simpson, A. P. 2003. Possible articulatory reasons for sex-specific differences in vowel duration. *Proc. 6th International Seminar on Speech Production* Sydney, 261–266.

[17] Sussman, H. M., Fruchter, D., Hilbert, J., Sirosh, J. 1998. Linear correlates in the speech signal: The orderly output constraint. *J. Behav. Brain Sci.* 21, 241–299.

[18] Sussman, H. M., Hoemeke, K. A., Ahmed, F. S. 1993. A cross-linguistic investigation of locus equations as a phonetic descriptor for place of articulation. *J. Acoust. Soc. Am.* 94, 1256–1268.

[19] Sussman, H. M., McCaffrey, H. A., Matthews, S. A. 1991. An investigation of locus equations as a source of relational invariance for stop place categorization. *J. Acoust. Soc. Am.* 90, 1309–1325.

[20] Traunmüller, H. 1988. Paralinguistic variation and invariance in the characteristic frequencies of vowels. *Phonetica* 45, 1–29.

[21] Traunmüller, H. 1990. Analytical expressions for the tonotopic sensory scale. *J. Acoust. Soc. Am.* 88, 97–100.

[22] Whiteside, S. P. 2001. Sex-specific fundamental and formant frequency patterns in a cross-sectional study. *J. Acoust. Soc. Am.* 110, 464–78.