A Case Study on Gender and the Acoustic Vowel Space

Summer Chambers – University of Virginia

LING 5409: Acoustic Phonetics

December 2019

This study explores differences in female and male production of English vowels. While many phoneticians have observed differences in the sizes of the female and male acoustic vowel space, the reasons for these differences are uncertain. Cardinal vowels and their formant frequencies are analyzed in three types of speech. Results show that differences in acoustic vowel space cannot be entirely explained by differences in fundamental frequency, and that speakers effectively manipulate vowel space size when imitating the speech of the opposite gender. These conclusions have implications for future research into the physical and social causes of gender differences in speech.

I. Introduction

After a long history of using male speech as a default in descriptive phonetics, many modern phoneticians have become interested in the differences between female and male speech. Several differences have been proposed in terms of vocal quality—such as breathy voice—and duration—voice onset time for stops or vowel length contrasts (Simpson, 2009). But, arguably, the most noticeable difference between female and male speech, and the one most accessible to non-linguists, is pitch. In phonetics, pitch is generally comparable to fundamental frequency, the frequency at which the vocal folds vibrate. Due to differences in the anatomy of female and male vocal folds (male vocal folds are thicker and longer), the average fundamental frequency is 200 Hz for females and 100 Hz for males (Simpson, 2009). This means the pitch of the female speaking voice is on average about an octave higher than the male voice.

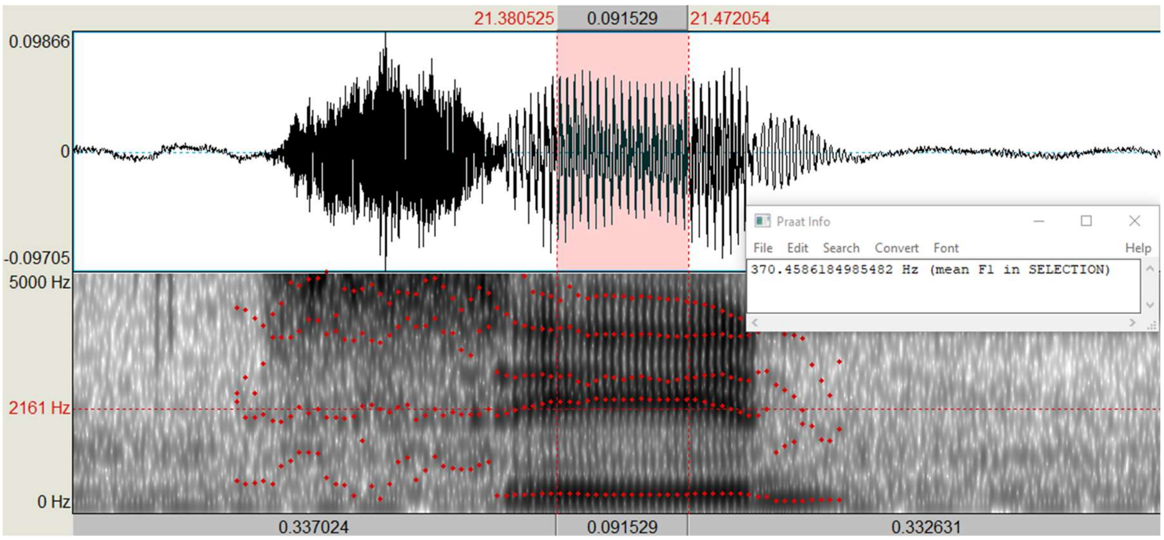
Relatedly, after puberty, the male larynx lowers down, causing a substantial difference in the length of the female and male vocal tract (the pharynx). The average length for females is about 14.5cm while the average length for males is around 18cm (Simpson, 2009). This difference is thought to be responsible for acoustic differences in the female and male voice, particularly in the resonant frequencies of the vocal tract that define different vowel qualities (formants). The values of the first two resonant frequencies—essentially the most amplified overtones of the voice—are called F1 and F2, and can be used to estimate the size of the acoustic vowel space. Despite the general consensus that women produce higher formant frequencies, it is often the male values that phoneticians use to define the vowels of any given language.

Previous experiments and formant analyses have shown the female acoustic vowel space is larger than the male acoustic vowel space in multiple languages (Simpson, 2009). Interestingly, the vowel spaces differ non-uniformly, that is, certain vowels have differences in F1 and F2 variation, and no one constant can be used to map the male space onto the female space or vice versa (Simpson, 2009). Many reasons have been proposed for why these non-uniform differences exist, but none has been validated. Diehl et al. (1996) suggest that the larger size of the female vowel space could be a direct result of the higher female fundamental frequency. To test their hypothesis, I sought to map out the size of the female and male vowel spaces and determine their relationship to fundamental frequency.

II. Methods

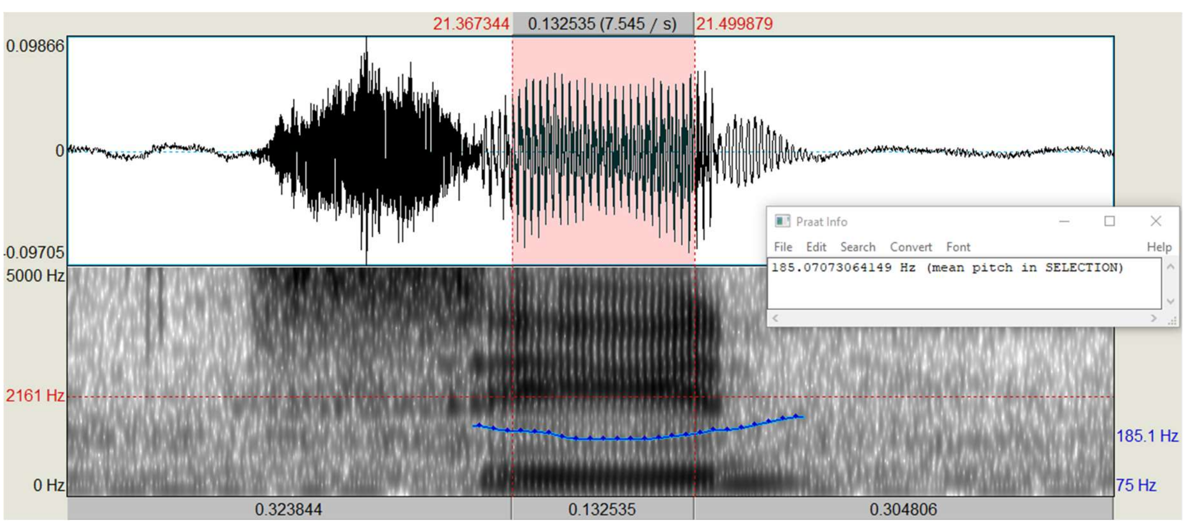
The subjects for this study included three women and three men between the ages of 18 and 24. Each subject was provided with a paper copy of the following list of seven monosyllabic English words: Seed, Dusk, Sad, Disk, Sued, Desk, Sod. The words structured [dVsk] were "distractors", while those structured [sV(V)d] were of interest to this study. These four vowels [i] [æ] [ɑ] [u] represent cardinal vowels, the four most extreme points of the acoustic vowel space. The list was read by each subject three times. The voice data were recorded in one long (~40s) .wav file for each subject, using a USB Blue

Snowball Tripod Microphone. After all the data were collected, the purpose and details of the experiment were explained to the subjects, who were compensated for their time. The data were analyzed primarily using Praat (Boersma and Weenink, 2019). For each word repetition, I selected a stable portion of the vowel and used Praat's "Formants" function to record both F1 and F2 as shown in Figure 2A.



2A: Formant Extraction with Praat

Next, I selected the fourth word in each of the subject's three readings and used Praat to estimate their average fundamental frequency, as shown in figure 2B.



2B: Pitch Extraction with Praat

Reading 1: Sung Vowels

For the first reading, subjects of either gender were instructed to sing the words on the provided pitch of Bb3, approximately 233Hz. In this way, the variable of fundamental frequency was eliminated. The pitch was played for each subject from the computer before they began, and all subjects successfully matched pitch.

Reading 2: Spoken Vowels

For the second reading, subjects were asked to read the words in their normal speaking voice. As expected, fundamental frequency was quite distinct for women and men, though it also varied within each gender group.

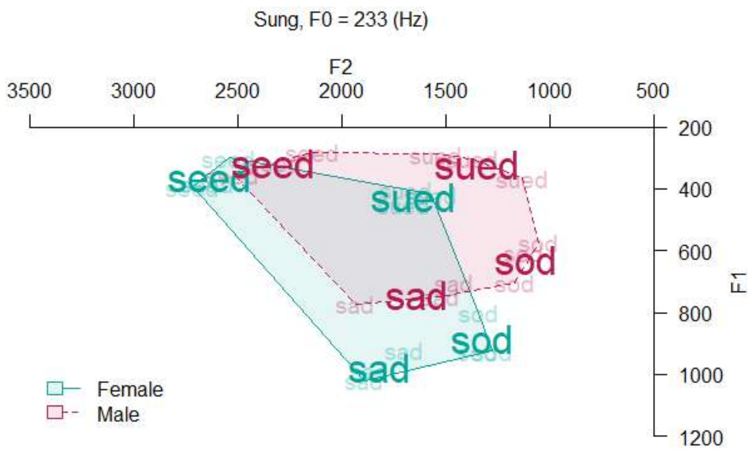
Reading 3: Imitated Vowels

For the third reading, the three female subjects were asked to read the words "like a boy" while the three male subjects were asked to read the words "like a girl". When formant and pitch values were all recorded in Microsoft Excel, I analyzed data using R (R Core Team 2014, McCloy 2016, Wickham, 2009).

III. Results

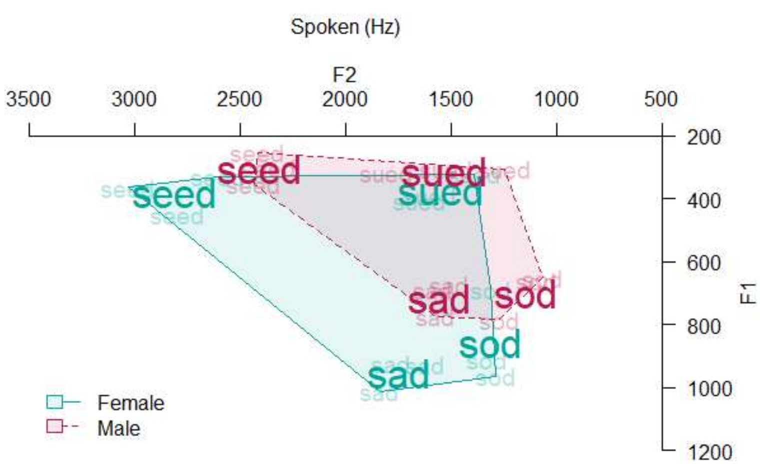
For each reading type, graphs were created by plotting F2 values on the X axis and F1 values on the Y axis. Both axes were reversed so that values run high to low rather than low to high. In this way, the graphs more closely resemble the articulatory vowel space. The left figure in the following pairs of graphs shows all data points for each reading, with shaded polygons connecting the most extreme values within each gender group and the gender means of each vowel shown in bold type. The right figure shows just the means of each gender and the shaded polygon outlining each gender's mean acoustic vowel space.

Figure 3A shows that even on the same pitch (eliminating F0 as a variable), there is a difference in the size and placement of female and male acoustic vowel space.



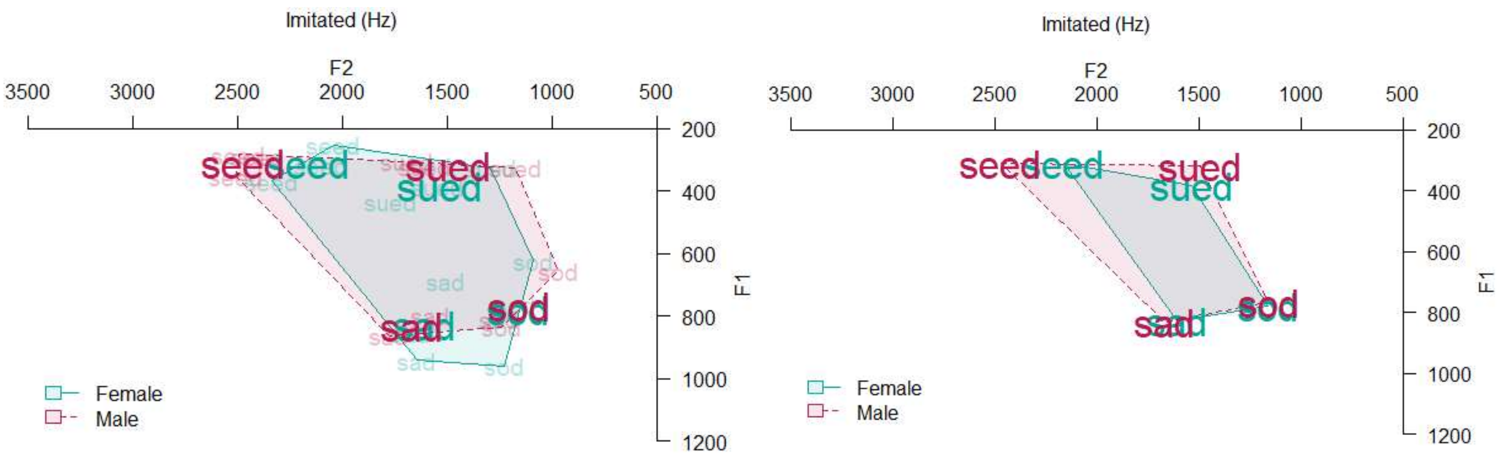
3A: Sung Vowels and Means by Gender

Figure 3B shows an even more extreme difference in vowel space. Again, notice that the differences in the female and male spaces are not uniform. The mid-low female vowel in the word /sad/ has a significantly higher F1 value (associated with openness) than its male equivalent, while the mean female high-front vowel in /seed/ is mostly distinguished by its higher F2 value (associated with frontness).



3B: Spoken Vowels and Means by Gender

Figure 3C shows that the female vowel space has shrunken inside the male space, which has expanded slightly.



3C: Imitated Vowels and Means by Gender

The area calculated from the gender-mean polygons is shown in table 3D.

Gender	Avg Area (kHz^2)	Reading
Female	403.4566	Sung
Male	300.0551	Sung
Female	481.5685	Spoken
Male	258.8452	Spoken
Female	238.3331	Imitated
Male	377.5232	Imitated

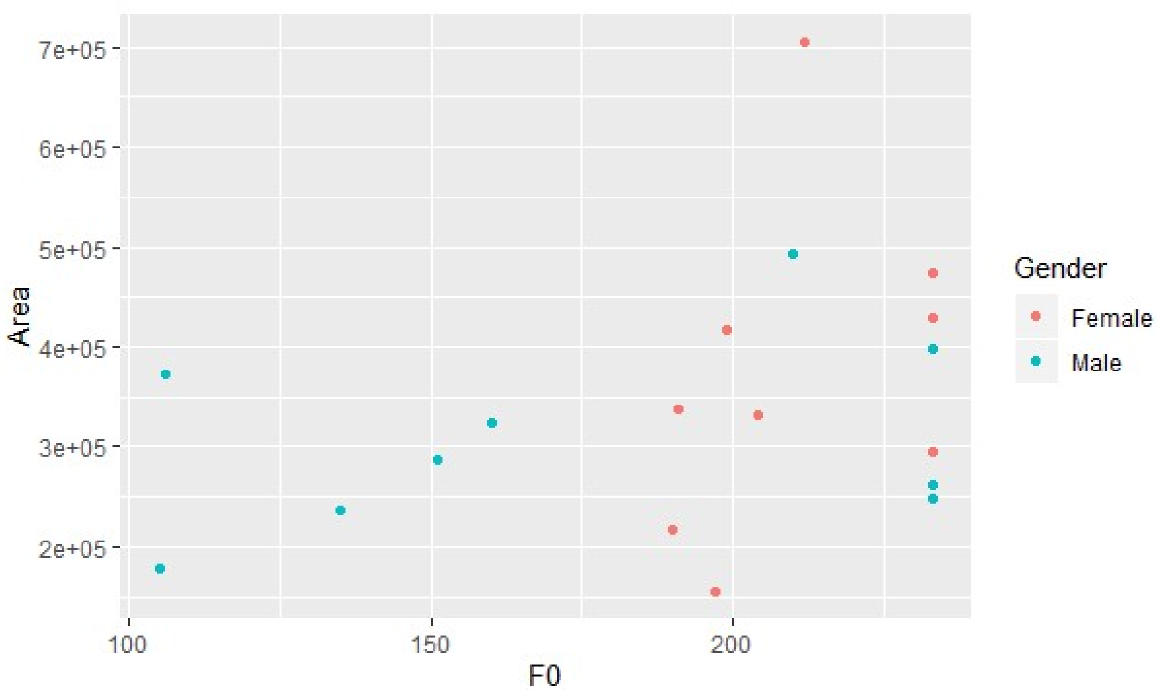
3D: Average Female and Male Vowel Space Area by Reading

Regarding the relationship between fundamental frequency and acoustic vowel space, we have the following data in table 3E:

Averages	Female	Male
Sung Area (kHz^2)	403.4566	300.0551
Sung F0 (Hz)	233	233
Spoken Area (kHz^2)	481.5685	258.8452
Spoken F0 (Hz)	205	115
Imitated Area (kHz^2)	238.3331	377.5232
Imitated F0 (Hz)	193	174

3E: Female and Male Average Vowel Space Area and Fundamental Frequency

For the sung reading, we see that the female vowel space is larger by more than 100 square kilohertz, even though fundamental frequency is constant. Clearly, the vowel space area cannot be simply computed from the fundamental frequency. For the spoken reading, the difference is even more dramatic, at more than 200 square kilohertz. Finally, we see that females imitating male speech significantly reduced their vowel space size so that it was even smaller than the spoken male space. Males imitating females increased the size of their vowel space, though not as dramatically as the females did, nor reaching the actual size of the spoken female space. This occurs despite the fact that the average female F0 is still above the average male F0, suggesting that the subjects are adjusting more than just their fundamental frequency when imitating the opposite gender.



3F: Fundamental Frequency (Hz) vs. Vowel Space Area (Hz^2)

The correlation between fundamental frequency and vowel space area is a weak, though vaguely linear *r = 0.33*.

IV. Discussion

From my experimental results, the female vowel space is on average larger than the male vowel space, in terms of acoustics. It should be noted that these observed differences do not tell us anything definite about the articulatory vowel space, that is, the physical movements of the tongue, teeth, jaw, etc. The data do show that this size differential cannot be totally explained by the higher female fundamental frequency. However, due to the more extreme difference in vowel space area for the spoken reading, fundamental frequency appears to have some influence over vowel space area, although the magnitude of the effect is unclear. We're now left wondering what exactly explains this size differential between females and males.

Some researchers have proposed that men might use the same vocal articulations they learned before puberty even after their pharynx has lengthened, which could cause the non-uniform differences in vowel space (Traunmüller, 1984). Others say that the smaller size of the female vocal tract might allow women to reach their articulatory targets in less time and therefore generate more extreme vowels (Simpson, 2009). Some believe that the female voice doesn't define vowel quality as well due to the logarithmic nature of the harmonic scale, so women space vowels apart farther to compensate for this (Diehl et al., 1996). While nothing is conclusive, most agree that the final explanation must be at least partly socio-phonetic—men and women learn to speak differently according to their gender (Simpson, 2009).

Much linguistic research into gendered speech suggests that women use more standard or prestigious forms of language, perhaps as a way to assert social position in a male-dominant society (Coates, 1986). It's not totally accepted that women are more easily understood because of the size of the vowel space, but some say that since women produce more open variants of vowels (signaled by higher F1 values), their speech is more phonetically explicit (Henton, 1995). Interestingly, while the acoustic size of the vowel space is not a particularly accessible feature of female and male voice, speakers did adjust this size in imitation of the opposite gender, whether consciously or unconsciously. In the past, it's been proposed that women are more in tune to linguistic differences than men, which would partially explain why women use more prestige forms (Chambers, 1995). This might suggest that women would be more successful at imitating male speech. Future studies could examine the articulatory movements made by females and males imitating the opposite gender, to see how the vocal tract is manipulated. It would also be valuable to consider how speakers perceive female and male speech qualitatively, in terms of clarity and precision.

References

Boersma, Paul & Weenink, David (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.08, retrieved 5 December 2019 from http://www.praat.org/

Chambers, J. K. (1995). *Sociolinguistic Theory*. Oxford: Blackwell Publishers Inc.

Coates, Jennifer (1986). *Women, Men and Language: A sociolinguistic account of sex differences in language*. London: Longman.

Diehl, Randy L., Bj¨orn Lindblom, Kathryn A. Hoemeke, and Richard P. Fahey (1996). On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics*, 24:187–208.

Henton, Caroline (1995). Cross-language variation in the vowels of female and male speakers. In *Proceedings of the XIIIth International Congress of Phonetic Sciences* (ICPhS 95), 420–423.

McCloy, Daniel R. (2016). phonR: tools for phoneticians and phonologists. R package version 1.0-7. http://drammock.github.io/phonR/

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Simpson, A. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, 3(2), 621-640. https://doi.org/10.1111/j.1749-818X.2009.00125.x

Traunmüller, H. (1984). Articulatory and perceptual factors controlling the age- and sex-conditioned variability in formant frequencies of vowels. *Speech Communication, 3*(1), 49–61. https://doi.org/10.1016/0167-6393(84)90008-6

Wickham, H. (2009) ggplot2: elegant graphics for data analysis. Springer New York.