Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Juliana Dieterich Boff March 30, 2018

Gender Differences in Acoustic Speech Perception and Production:

Effects of articulatory flexibility on performance and intelligibility of speech

by

Juliana Dieterich Boff

Dr. Lynne Nygaard, Ph.D.
Adviser

Department of Psychology

Lynne Nygaard, Ph.D.

Adviser

Harold Gouzoules, Ph.D.

Committee Member

Susan Tamasi, Ph.D.

Committee Member

2018

Gender Differences in Acoustic Speech Perception and Production:

Effects of articulatory flexibility on performance and intelligibility of speech

By

Juliana Dieterich Boff

Lynne Nygaard, Ph.D.

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Psychology

2018

Abstract

Gender Differences in Acoustic Speech Perception and Production:

Effects of articulatory flexibility on performance and intelligibility of speech

By Juliana Dieterich Boff

Speech production is characterized by a considerable amount of individual and group variation. The current study investigated pronunciation differences between male and female speakers, as well as how a lexical characteristic such as word frequency affected productions. I hypothesized that males and females would differ in the production of speech due to *articulatory flexibility*, or how distinctively the acoustic-phonetic properties of each utterance were produced. In this study, the effects of speaker gender (male, female) and word frequency (high, low) on fundamental frequency (F0), word and vowel duration, and by-vowel and by-speaker vowel dispersion were examined. Female speakers were hypothesized to show greater articulatory flexibility and to change their productions more as a function of word frequency than males. Although measures of vowel dispersion did not differ significantly as a function of speaker gender or word frequency, there were significant effects of both word frequency and gender on duration measures. Both vowel and word durations were longer for low frequency words, but females varied their vowel duration more as a function of word frequency than did males. These differences in duration might be related to articulatory effort, and indicative of articulation flexibility. These differences could be one way in which individuals "perform" their gender in social communication. Future research should consider expanding on the idea that these differences in pronunciation are meant to perform gender and include a perceptual analysis of speech to observe any effects of gender and articulatory flexibility on intelligibility.

Gender Differences in Acoustic Speech Perception and Production:

Effects of articulatory flexibility on performance and intelligibility of speech

By

Juliana Dieterich Boff

Lynne Nygaard, Ph.D.

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Psychology

2018

Acknowledgements

This research was funded by an Independent Research Grant from Emory University. I thank Dr. Lynne Nygaard and Dr. Eva Lewandowski for taking on this project with me. I couldn't have done it without their mentoring and support. I would also like to thank my parents, for their love and encouragement every step of the way.

Table of Contents

Gender Differences in Speech Production:
Effects of Articulatory Flexibility on the Acoustic Structure of Speech

From differences in the basic perception and production of speech to differences in higher level aspects of language such as word choice and syntactic processing, spoken language is characterized by a considerable amount of individual variation (Chiffi, 2012; King & Just, 1991; Miller, 2011; Pollack, Pickett, & Sumby, 1954). One particularly salient characteristic of speakers that may account for some of this variation is the gender of the language user. Gender[1] is a prominent component of spoken communication, and humans have little difficulty identifying the gender of a talker from auditory information alone (Li & Fu, 2011; Mullennix, Johnson, Topcu-Durgun, & Farnsworth, 1995; Nygaard & Queen, 2000). Previous research has shown that males and females may use different communicative strategies. For example, males and females appear to differ in their choice of words and conversational topics (Eckert, 1989; Haas, 1979; Lakoff, 2004). Analyses of spoken language reveal that females use more references to feelings and more evaluative and interpretive terms, while males refer more often to time and physical movement (Haas, 1979). Differences have also been observed in topic and pronoun use in written language. Newman, Groom, Handelman, and Pennebaker (2008) found that females were more likely to refer to psychological processes than males and males included more references to object properties than females. These results provide evidence for gender differences in speech communication. The goal of the present work is to examine the impact of linguistic characteristics on gender differences in speech production.

[1] In the present study, "gender" is defined as the participants' self-reported category as either male or female. I understand this categorization doesn't fully capture either gender identity or biological sex. The relationship between sex and gender complex, and not easily defined using objective measures. The present definition is thought to capture enough information to interpret initial findings on gender performativity, but further elaboration is needed in future research.

Gender Differences in Communication

Differences between males and females are not unique to spoken language and emerge even in non-linguistic vocal behaviors. For example, differences have been found between males and females in the timing and quality of laughter (Provine, 1993). The study of laughter is informative because it has been recorded in other primate species, such as chimpanzees, suggesting that this type of social, non-linguistic vocalization may be evolutionarily preserved (Provine, 2017). Despite the non-linguistic nature of the vocalization, laughter appears to be used as a social instrument and to differ across male and female interlocutors (Bilous & Krauss, 1988). Provine (1993) found that females laughed more than males in conversations between heterosexual pairs. Bilous and Krauss (1988) propose that males and females differ on laughter due to desire to elicit positive evaluations from members of the opposite sex. These findings suggest that there are differences between males and females in social communication, both in language and in other types of vocalizations. Speech and non-speech behaviors may differ in evolutionary origin and amounts of cognitive effort involved, suggesting gender differences are present in domains with varying degrees of conscious effort demands (Sweller & Jose, 2011). Such evidence further suggests that differences between males and females may be highly ingrained or outside of effortful control.

In addition to differences in the *pragmatics* of language use, males and females differ in the physical *production* of speech. One of the most significant differences between male and female voices is their fundamental frequency (F0) or pitch: females typically have higher F0 than males. The pitch variations between males and females have been attributed to differences in the size and shape of the vocal tract that vary reliably across the sexes (Fitch & Giedd, 1999; Titze, 1989). That is, males have longer vocal tracts, which produce slower and deeper pitch

resonances whereas females have shorter vocal tracts, which produce faster and higher pitch resonances. Analyses of anatomical differences in male and female vocal tracts, such as Fitch and Giedd's, highlight a physical and biological distinction that may account for some of the gender differences observed in the production of speech, particularly differences in pitch. Although there are well-studied differences between males and females in the anatomy and physiology of the vocal apparatus, these differences do not fully account for gender variation in the production of speech (Fraccaro et al., 2013; Strand & Johnson, 1996).

Research has shown that males and females differ in the way they produce speech for reasons other than anatomical differences (i.e., the size and shape of the vocal tract; (Diehl, Lindblom, Hoemeke, & Fahey, 1996; Johnson, 2006; Stuart-Smith, 2007), and that these differences have social consequences. Fraccaro et al. (2013) showed that differences in F0 affected perceived attractiveness of the speaker and even perceived dominance in the interaction. Both females and males found "prototypical" speakers of the opposite gender (e.g., higher pitch for females, and lower pitch for males) to be more attractive. More interestingly, speakers instructed to over-exaggerate their pitch to sound more atypical according to their sex, were rated by their counterparts to be less attractive. The effect was not significant when exaggerations were made to sound more typical of one's sex. These findings show that pitch is an easily identified characteristic of male and female speech, which in turn can affect the perceptions and attitudes of language users. Furthermore, habitual F0 range may be defined by anatomy but is nonetheless somewhat controllable by the language user. Gender differences in speech production, easily perceptible in conversation, have important implications for social interaction and how we form impressions of others.

Individual and Group Differences in Pronunciation

Each individual has a unique way of speaking, a unique voice accompanied by unique production characteristics, which is noticeable in everyday conversations. Research shows that, as with identifying speaker gender, people can easily ascertain speaker identity from vocal information alone (Pollack et al., 1954). These talker-specific characteristics, or idiolects, are well studied in the context of spoken communication and speech perception (Chiffi, 2012; Miller, 2011) and the variations in individuals' speech have been shown to come from sources beyond vocal tract anatomy. Kraljic, Brennan, and Samuel (2008) propose the main source of variation in production comes in fact from each individual's pattern of articulation, which is rehearsed or learned throughout our lifetime. These authors highlight the importance of non-biological differences that account for differences in speech production.

Similarly, the notion of regional dialects, or distinct ways of speaking a language depending on one's location or cultural group, is evident in common social interactions. Such variations manifest in pronunciation, word choice and spelling across different countries speaking a same language, or even in two different communities in the same city (Macaulay, 2010). Johnson (2006) shows vowel formants, specifically F1 and F2, differ according to language and dialect of a same language, even when accounting for body and vocal tract size differences across populations and cultures, suggesting the variations are not explained solely by anatomical differences. Macaulay (2010) proposes that dialects indicate differences in prestige and social status between groups in a same location, supporting the idea that dialects are socially constructed. Perhaps the most intriguing evidence comes from a recent study by Sneller and Roberts (2018), where they investigated why some dialects get incorporated into social groups, while others are not used and disappear. The hypothesis was that desirable traits of a group

would affect whether their speech would be adopted by others. In their experiment, using a made-up language, a group identified as having the desirable quality of "toughness" did in fact have their dialect adopted by other participants. Importantly, when the trait was not considered socially relevant, the effect disappeared. Sneller and Roberts conclude that dialect propagation is in some way "biased", and individuals take social cues into consideration when adopting (or failing to adopt) dialectal variations. Kraljic et al. (2008) propose that while individual variations could arise from motor articulation and implementation habits, the source of dialectal variation could may be related to changes at the phonological level. If that is the case, dialect variations in speech sounds should be more malleable. However, learning would require exposure to meaningful language such as words, as opposed to single sounds or non-word utterances. The authors argue that the motor articulation differences they describe as the source of idiolect variation would not be a viable alternative for dialects, and that excessive training would be required in order to make a dialect consistent across groups of people.

I hypothesize that differences in speech as a function of gender are similar to dialect or idiolect variations. Since differences between males and females cannot be fully accounted for by anatomical differences, I propose that social factors may influence speech, leading males and females to use their productions as a way to perform or construct gender, in a way similar to how language is used for the same purpose (Eckert & McConnell-Ginet, 1992). For example, males and females can differ in the ways in which particular speech sounds are produced. Strand and Johnson (1996) found that males and females differ in how the fricative /s/ is produced, such that the boundary between the /s/ and /ʃ/ fricative categories occurs at a lower frequency for males than for females. When investigating sexuality in a conservative community in California, Podesva and Hofwegen (2015) found that these variations are greater in communities with strong

heteronormative values, suggesting that the acoustic-phonetic properties of language can be leveraged to convey gender. Stuart-Smith (2007) supports the idea that the variation in the production of the /s/ sound is a way in which language users "construct gender", as well as to perform according to their social status. In her study of males and females in different socioeconomic groups in Glasgow, she noticed males' and females' productions of /s/ were different from each other for most social distinctions, but the productions from "working class" women were more similar to those of their male counterparts. This suggests an influence of social group as well as gender, favoring the view that these differences are used to perform gender as well as status. These sociolinguistic findings, together with the evidence stating that males and females vary in the ways they produce specific speech sounds even when accounting for anatomical differences (Johnson, 2005; Strand & Johnson, 1996; Stuart-Smith, 2007), support the view that speech production might be a way to "perform" gender, assuming that there is no clear and direct relationship between this performance difference and biological and anatomical determinants.

Gender Differences in Articulatory Flexibility and Precision

Males and females differ in the way they produce specific sounds, but variations are also present in how precise males and females are in their pronunciations overall. Females have greater *vowel dispersion*, a measure of how distinctly each vowel sound is produced. Bradlow, Torretta, and Pisoni (1996) examined the vocal properties of 20 speakers (10 male, 10 female) and found that female speakers had greater vowel dispersion, meaning their utterances of vowel sounds were more precise, or further away from the central point of their vowel space. While investigating acoustic properties according to speakers' sexual orientation, (Munson & Solomon,

2004) found similar results, with females having higher vowel dispersion across all sexual orientations, and significant differences in vowel space expansion as a function of sex. Diehl and colleagues (Diehl et al., 1996) argue that such production differences according to gender are not a product of vocal tract anatomy, since there is not a linear and "uniform" relationship between vocal tract length and vowel dispersion, meaning a talker's dispersion cannot be accurately predicted based on anatomical data such as body size and vocal tract length. Other research involving heterosexual and homosexual participants from both sexes has shown different results, where homosexual females had significantly lower vowel dispersion than their heterosexual counterparts (Rendall, Vasey, & McKenzie, 2008). Rendall et al. propose this dispersion difference according to orientation may come in part from differences in lip protrusion, which is a way to indicate assertiveness or dominance in interaction. On this view, lip protrusion could be another way these individuals perform their orientation.

Bradlow et al.'s (1996) also found that vowel dispersion of male and female speech had consequences for listeners' speech perception. The authors showed that female speakers were significantly more intelligible than males, with the top five most intelligible talkers in their sample being female. Furthermore, greater vowel dispersion was correlated with greater intelligibility, suggesting that speakers' ability to produce clear variations across speech sounds may be related to listeners' ability to understand words. This finding suggests that males and females may differ in their sensitivity to listeners' comprehension needs, as well as how these differences may be relevant for aspects of speech perception and speech communication more generally.

Consequences of Gender Differences in Speech and Language

As previous studies show, differences in the production of speech may play an important role in social interaction (Fraccaro et al., 2013; Macaulay, 2010; Newman et al., 2008; Owren & Bachorowski, 2003; Strand & Johnson, 1996). Another way in which gender differences can influence social communication has been reported by studies investigating vocal accommodation, a behavior in which individuals entrain to the speech characteristics of others. Some of these investigations indicate that females converge more than males; that is, females altered their productions more to match those of others than did males (Namy, Nygaard, & Sauerteig, 2002; Pardo, Jay, & Krauss, 2010; Shin, 2013). Namy et al. (2002) used listeners' judgments to determine whether a participant's shadowed utterances were more similar to the model talker's, or to the participant's original non-shadowed utterances. The listeners judged shadowed utterances from females to be more similar to the model talker's than those of males, suggesting that females accommodated more than males to the model talkers. The researchers propose such variation emerges from differences in females' sensitivity to perceptual features of speech, such as emotional tone of voice. The authors suggest that females may be socialized to be more attuned to social information carried by the voice, leading to greater assimilation in speech. Using a similar procedure with conversational phrases instead of single word utterances, Pardo et al. (2010) also found the amount of vocal accommodation was influenced by the gender of the talkers. Assuming that gender differences do play a role in vocal accommodation, one hypothesis is that since alignment requires changing one's productions, individuals with greater *articulatory flexibility* (i.e., more precise phoneme articulations as measured by-vowel dispersion) would more easily be able to match other speakers' productions and thus converge. If

females do in fact converge more than males, this could be because they are more flexible articulators initially, and are able to produce a greater range of more distinct phoneme utterances.

Gender differences in speech production and its influence on social interaction also extends to language change in society. Specifically, gender differences are observed with respect to setting societal and cultural change in motion. Sociolinguistic work suggests that females are the primary drivers of language change (Eckert, 1989; Shin, 2013). For example, Shin (2013) found native Spanish-speaking females raised in Latin America, not their male counterparts, drove changes in pronoun use in their communities in America. Females may be leading change due to multiple factors, and Shin hypothesized females more heavily impact language because they spend more time with American-born, English and Spanish bilingual individuals (i.e., their children), but also due to females' heightened sensitivity to variations in language and their use of it as a social instrument. Latino women in America changed their pronunciations more when compared to Latino men, and may have had more opportunities to propagate these changes in pronoun use. Thus, differences in the flexibility of articulation of speech sounds may contribute to the role of females in the process of language change. A more flexible articulator would produce phoneme sounds more distinctively from one another and have a wider range of productions to use when changing their speech. Furthermore, the clear distinction between phoneme utterances could be highlighting linguistic variations, which in turn would lead to broader dissemination of the variations. On this view, variants that are more intelligible or salient would in turn be more easily conveyed between individuals.

Linguistic Properties and Gender Differences in Pronunciation

One important aspect influencing gender differences seen both in vocal accommodation and language changes over time may be sensitivity to the distinctions within language itself. For instance, words used in conversation only occasionally, or low frequency words, are more precisely articulated to facilitate understanding, when compared to words used more often, or high frequency words (Pardo, Jordan, Mallari, Scanlon, & Lewandowski, 2013). Pardo et al. (2013) investigated how much males and females vocally accommodated to a model talker based on a linguistic item's usage frequency. Using perceptual judgments of similarity, the authors found that females were more likely to converge towards low frequency words than males, meaning females modified their speech more to match less commonly heard words more than males. I hypothesize that this effect could be due to females' larger articulation range (e.g., Bradlow et al., 1996), which allows them to vary their productions more broadly. While low frequency words must be more precisely articulated to be comprehended, females may have a greater ability to be change their pronunciations on average than males, leading to differences in production and thus greater convergence when saying these words.

Further research supports that females are more likely than males to change their vocal accommodation behavior depending on the lexical category of the stimulus word (Pardo, Urmanche, Wilman, & Wiener, 2017a). Pardo et al. (2017) replicated findings by Munson and Solomon (2004) that indicated low frequency words had higher vowel dispersion than high frequency words and further demonstrated that females vocally accommodated more to low frequency words than did males. From this evidence, it follows that that seldom used words are articulated more precisely than commonly used ones, as indexed by an expanded vowel space (i.e., high vowel dispersion). The results also highlight a difference in speech production based

on language characteristics that varies as a function of gender. The presence of differences in the production of speech according to word frequency across genders suggests that there may be differences in the articulatory flexibility of males and females for high and low frequency words. The proposal here is that, similarly to the effect observed with vowel dispersion, low frequency words would have higher articulatory flexibility rates than high frequency words, and females would show a greater difference when producing low frequency words relative to high frequency words than would males.

The Current Study

Although there is evidence suggesting differences in production may be related to both differences in lexical characteristics, such as word frequency, and individual characteristics, such as gender, there has not yet been a systematic investigation of the relationship between articulatory flexibility, word frequency, and differences in male and female speech. The current study investigates this topic through an analysis of speech production, using three measures of articulatory flexibility, to determine if males and females differ in terms of articulatory flexibility. The ability to change pronunciation to increase precision will be assessed both by vowel dispersion and overall articulatory variability in pitch and duration and by the extent to which pronunciation changes as a function of lexical characteristics such as word frequency.

The goal of this experiment was to investigate potential gender differences among several dimensions of speech. Male and female participants were recorded reading a list of words that varied in frequency of occurrence. These speech samples were then analyzed to examine potential differences across genders in fundamental frequency (pitch), word duration (speaking rate), vowel duration, and vowel dispersion. Following previous studies (Bradlow et al., 1996;

Fitch & Giedd, 1999; Fraccaro et al., 2013; Titze, 1989), I predicted the following: (1) females would have higher fundamental frequencies and higher vowel dispersion than males; (2) both word and vowel duration would be similar between males and females; and (3) pronunciation would differ as a function of word frequency, with higher vowel dispersion and longer durations for low than for high frequency words. Pitch was not expected to differ across word type. Finally, I examined whether males and females differed in their sensitivity to word frequency. I expected also that (4) female speakers would show greater sensitivity to lexical characteristics and therefore change their pronunciations to a greater extent across word type than would males. The data were analyzed as a function of talker gender, as well as stimulus word frequency.

Method

Speech stimuli. The speech samples used in this study were collected for previous work by the Speech and Language Perception Lab at Emory University (Lewandowski, 2018). The entire database included speech recordings from 44 native speakers of American English (10 male, 34 female) between the ages of 18 and 25. One male participant was excluded due to self-reported history of hearing or speech disorders. Speech samples from the remaining 9 males were included in analyses. To choose a subset of 9 female talkers, female talkers were matched to the male talkers according to their responses on a word familiarity scale that indicated participants' experience with the word stimuli (word details in the next section). Participants indicated their familiarity with each word on a 1 to 7 Likert-type scale (1 = do not recognize the word, do not know its definition; 4 = recognize the word, unsure of its definition; 7 = definitely recognize the word, definitely know its definition). In cases where multiple females matched a male's word familiarity scores, talkers were chosen randomly from the list of possible

candidates. In cases where no female matched a male's word familiarity score exactly, talkers were chosen randomly from a list including the females with the immediately higher and lower scores. This was done in an attempt to balance talkers' word familiarity scores in both groups, and to control for any articulation effects related to lack of understanding of the stimulus words. On average, the 18 talkers chosen for this study indicated high familiarity with both the high frequency ($\bar{x}$=6.95) and low frequency words ($\bar{x}$=6.58), although participants were unfamiliar with more of the low frequency words (2.60% rated 3 or lower) than high frequency words (.22% rated 3 or lower).

Word stimuli. Each speaker was recorded producing 96 bisyllabic words selected from *The English Lexicon Project* inventory (Balota et al., 2007). The word list was balanced for word frequency (Lund & Burgess, 1996), syllable stress, stressed syllable onset, and stressed syllable nucleus. Word frequency was characterized categorically as high or low. The high frequency words had a median frequency of 37,380.5 occurrences per 131 million ($\bar{x}$=72,945.33), while the low frequency words had a median frequency of 215.5 occurrences per 131 million ($\bar{x}$=620.23; Burgess & Livesay, 1998). The following vowel categories were represented in the stressed syllable nuclei: 23 /i/ , 24 /ɛ/, 25 /ɑ/, 24 /u/. Sound Studio sound editing software (Version 4, Felt Tip Inc., 2017) was used to amplitude normalize the files at 59.70% (-4.48 dB), and resample at 22,050 Hz. Sounds were segmented into individual files for each word utterance for each speaker, removing portions of noise and silence in between utterances.

Acoustic measures. Assessments of fundamental frequency (F0), word duration, vowel duration, and vowel dispersion were conducted to investigate gender differences in speech production. All measures were obtained using custom Praat scripts (Boersma & Weenink, 2017). F0 was measured to replicate a vast body of work supporting the existence of a gender difference

in this dimension (Babel & Bulatov, 2012; Fraccaro et al., 2013). Word and vowel durations were expected to be similar across males and females (Bradlow et al., 1996). Vowel space dispersion was assessed both on a by-vowel and by-speaker basis, as in Bradlow et al., (1996). The differences between these two types of dispersion are illustrated in Figure 1. The dispersion by-vowel allowed me to observe how consistently the speaker targeted his or her productions within a single vowel category. This was obtained by calculating the mean first two formants (F1 and F2) from all the productions of a vowel and then taking the Euclidean distance between each token and the vowel's center.

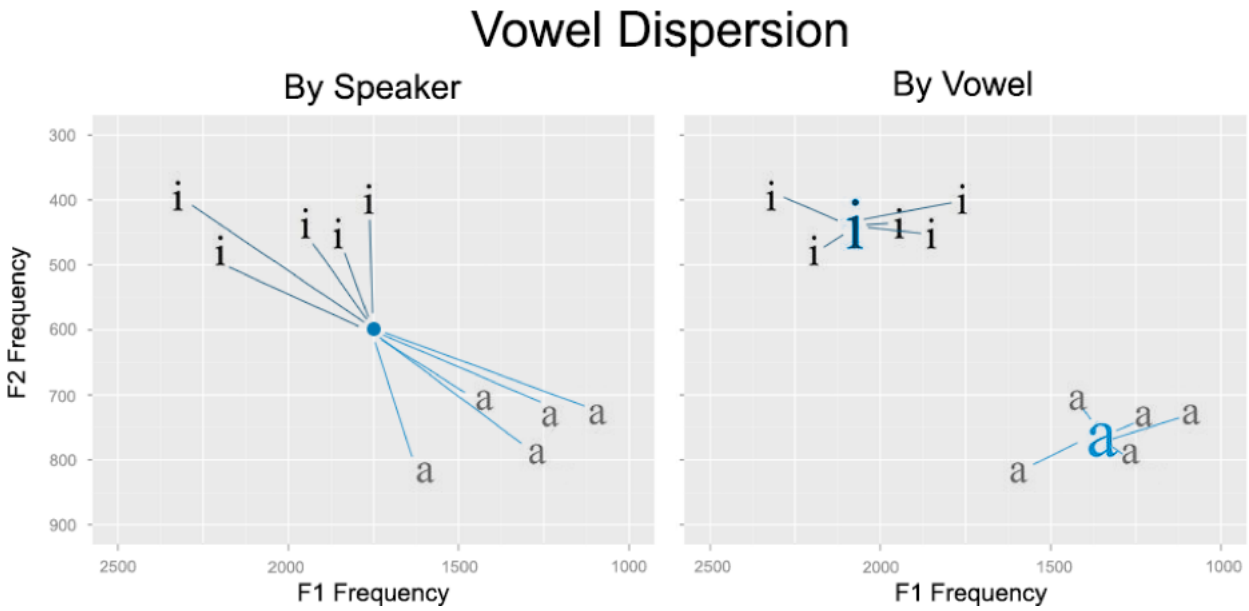

Figure 1. Illustration of vowel dispersion by-speaker and by-vowel. Reproduced with permission from Lewandowski (2018).

The dispersion by-speaker gave me the opportunity to examine global articulatory precision of each vowel sound compared to the center of that speaker's vowel space. By-speaker dispersion was calculated by taking the Euclidean distance between each vowel's F1 and F2 from the average F1 and F2 of all the utterances produced by that speaker). Custom Praat and R scripts (Boersma & Weenink, 2017; RCoreTeam, 2016) were used to norm the vowels using the

Lobanov technique (Lobanov, 1971), to merge the files for analysis, and to calculate the Euclidian distances, both according to speaker and vowel. The Lobanov technique was chosen because it has extensive empirical support and has been used in previous studies with similar aims (Babel, 2012). While the Lobanov normalization is used to preserve sociolinguistic variation from speech, controlling for anatomical and physiological differences among speakers (Adank, Smits, & Hout, 2004; Fabricius, Watt, & Johnson, 2009; Flynn & Foulkes, 2011), it might have also removed some of the speaker's individual characteristics (see Discussion).

Results

The data were collapsed across items and then submitted to a two-way mixed analysis of variance (ANOVA) with gender (male, female) as a between-subjects factor and word frequency (high, low) as a within-subjects factor. The dependent measures of F0, word and vowel duration, and vowel dispersion by-speaker and by-vowel were investigated using separate ANOVAs. For any analysis with statistically significant interaction effects ($p<.05$), post-hoc t-tests were conducted to obtain further information about relationships among the variables.

Fundamental frequency. Figure 2 shows the results for vocal pitch. As expected, there was a main effect of speaker gender on pitch, $F(1,16)=64.80$, $p<.001$. The direction of the effect indicates that females had generally higher pitch than males. Word frequency and the word frequency x gender interaction were non-significant, both $F(1,16)<0.677$, both $p>0.423$, indicating that pitch did not differ across word types and that speakers did not differentially use pitch for high and low frequency words.

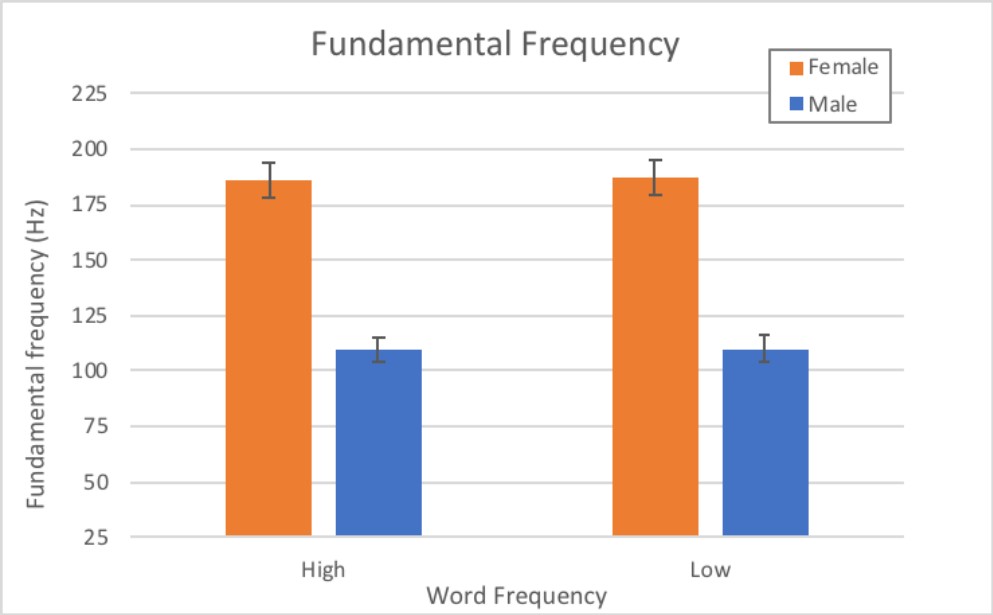

Figure 2. Fundamental frequency as a function of gender and word frequency. Error bars indicate standard error of the mean.

Word duration. Figure 3 illustrates the association between gender and word frequency on word duration. No main effect of gender was observed on word duration, $F(1,16)=2.99$, $p=0.103$. However, there was a reliable effect of word frequency, $F(1,16)=59$, $p<.001$, such that speakers generally used slower speaking rates for low frequency words. The word frequency x gender interaction was marginal, $F(1,16)=3.7$, $p=0.0723$. Post-hoc comparisons indicated that male word durations were marginally shorter (Bonferroni-corrected $\alpha=.025$) than female word durations for high frequency words ($t(16)=-2.14$, $p=0.0485$), but that word duration did not differ significantly for males and females on low frequency words ($t(16)=-1.37$, $p=0.189$). Both males and females used significantly longer word durations for low frequency words (females, $t(8)=-3.335$, $p=0.021$; males, $t(8)=-9.509$, $p<0.001$).

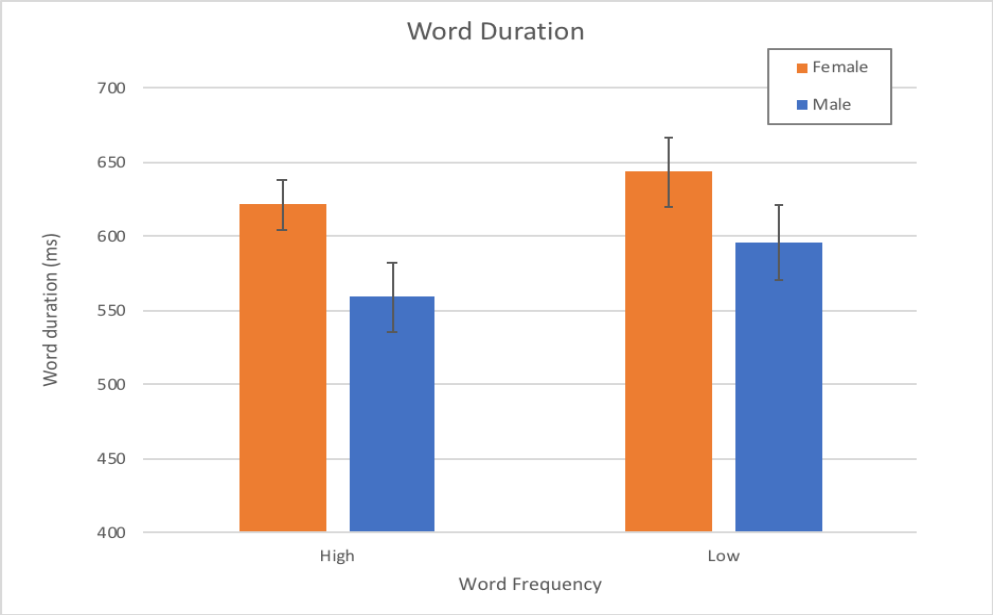

Figure 3. Word duration as a function of gender and word frequency. Error bars reflect standard error of the mean.

Vowel duration. While word duration provides a proxy for global speaking rate, vowel duration is an index of the emphasis placed on the vocalic portion of a word. The effects of gender and word frequency on vowel duration are depicted in Figure 4. Reliable main effects of both gender ($F(1,16)=15.3$, $p=0.001$) and word frequency ($F(1,16)=19.6$, $p<.001$) were observed, indicating females had longer vowel durations than males, and that vowel duration was longer for low frequency words. However, these main effects were conditioned by a significant word frequency x gender interaction, $F(1,16)=7.53$, $p=0.0144$. Post-hoc comparisons indicated that females used significantly longer vowel durations for low frequency words, $t(8)=-4.42$, $p=0.00223$, but that males' vowel durations did not differ by word frequency, $t(8)=-1.44$, $p=0.189$. Female vowel durations were significantly longer than male vowel durations for both high ($t(16)=-3.19$, $p=0.0057$) and low frequency words, $t(16)=-4.49$, $p=0.00037$.

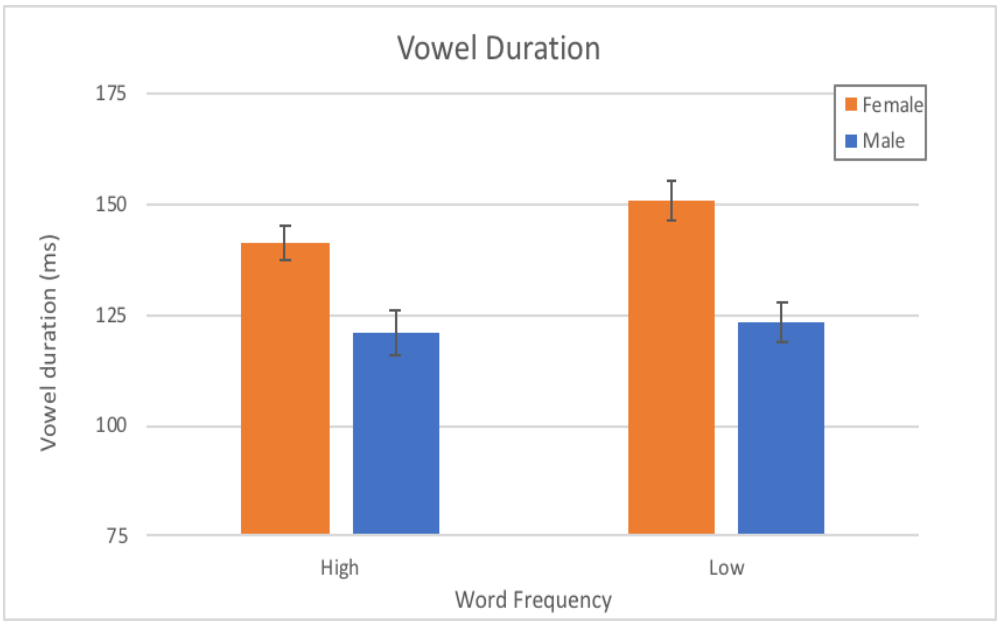

Figure 4. Vowel duration as a function of gender and word frequency. Error bars reflect standard error of the mean.

By-speaker vowel dispersion. By-speaker vowel dispersion is shown in Figure 5. No reliable effect of either factor of interest (speaker gender, word frequency) was observed on by-speaker vowel dispersion, all $F<2.28$ all $p>0.151$. This means dispersion from the speaker's vowel center did not significantly vary according to speaker gender or word frequency.

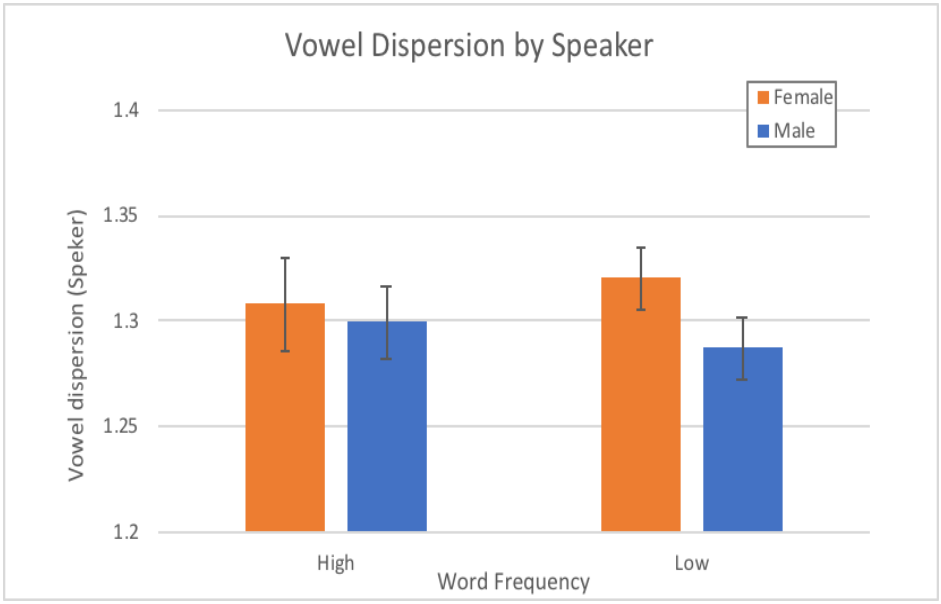

Figure 5. Vowel dispersion by-speaker, as a function of gender and word frequency Error bars reflect standard error of the mean.

By-vowel vowel dispersion. The effect of gender and word frequency on by-vowel dispersion is presented in Figure 6. No main effect of gender was observed on by-vowel dispersion, $F(1,16)=0.195$, $p=0.665$. A marginal main effect of word frequency, $F(1,16)=3.72$, $p=0.0717$ was observed, indicating slightly higher by-vowel vowel dispersion for low frequency words relative to high frequency words, meaning vowels in low frequency words were being pronounced with more variation or dispersion than vowels in high frequency words. The word frequency x gender interaction term was not statistically reliable, $F(1,16)=2.41$, $p=0.14$.

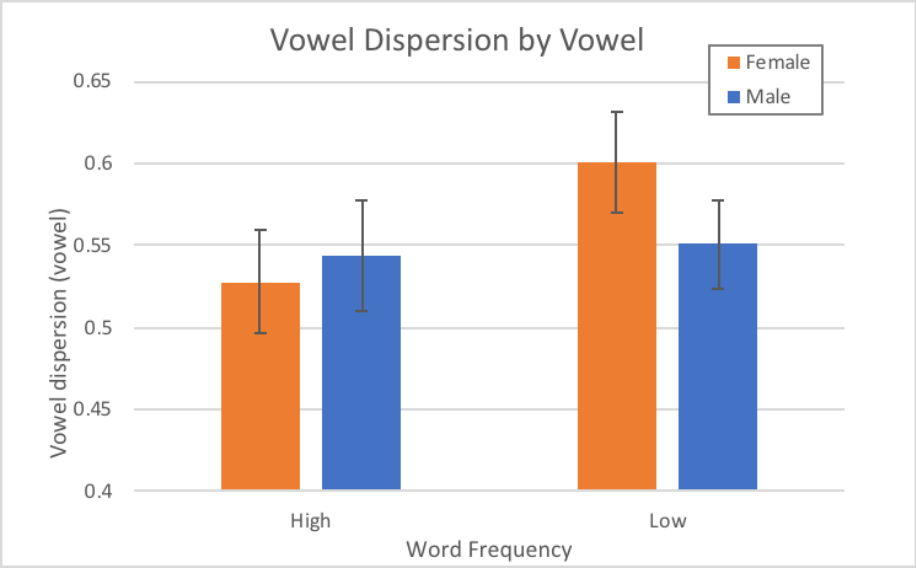

Figure 6. Vowel dispersion by-vowel, as a function of gender and word frequency. Error bars reflect standard error of the mean.

Discussion

This study investigated the effect of gender and word frequency on speech production, particularly by examining measures of articulatory flexibility. Findings showed that females had higher fundamental frequency than males, as expected from previous research (e.g., Fraccaro et al., 2013). In addition, I also observed unexpected differences in word and vowel duration. For

vowel duration, I found both an overall difference between males and females and a difference in the degree to which word frequency influenced vowel duration. Females placed significantly more emphasis on vowels than males, and their vowel durations were significantly longer for both high and low frequency words. Females also seemed to place more emphasis on the vocalic portions of low frequency words compared to high frequency words than did males. For word duration (i.e., speaking rate), there was a significant effect of word frequency, which indicated low frequency words were produced with longer duration, or slower speaking rates, than high frequency words. Although there was no significant effect of gender on word duration, males had marginally faster speaking rates for high frequency words than females. Finally, no significant differences in vowel dispersion were found, either as measured by-speaker or by-vowel. However, by-vowel vowel dispersion was marginally higher for low than for high frequency words.

Based on previous studies investigating individual differences and pronunciation changes in speaking rate and vowel duration (eg. Bradlow et al., 1996), I did not expect to see differences in speaking rate and vowel duration based on either word frequency or gender. I hypothesize that these differences might be representative of the fact that low frequency words are articulated more carefully in order to be understood (Pardo et al., 2013). Slower productions for low frequency words might facilitate their comprehension, given they are less familiar to the speakers. In our sample, females not only produced longer vowel durations than males but they also produced longer vowel durations for low as compared to high frequency words, suggesting that females may be more flexible in their production of vowels and perhaps more sensitive to the lexical characteristics of words. Bradlow et al. (1996) observed differences in duration of the /s/ sound when compared to the preceding and following syllables. They observed that females

had shorter durations for the sound, but in this case note that the correct (and more precise) production of the /s/ sound is in fact the shorter variant. Therefore, the researchers argue this difference in duration reflects greater precision in articulation by the female participants.

Vowel duration could also be reflecting increased effort in speech production. Perkell, Zandipour, Matthies, and Lane (2002) show clear, or more intelligible speech, is more effortful to produce than conversational speech and vowel duration accounted for up to 89-96% percent of the variance in the intelligibility across the study's participants. If larger vowel duration in our study reflects more effortful (and thus clearer) speech, it can also be considered an indicator of greater articulation flexibility. Furthermore, it would also explain some of the variation according to word frequency; it is understandable that low frequency words would need more effort in production, since they are encountered and/or said less often. In our sample, females displayed duration differences across word frequencies, while males did not. (Bradlow et al., 1996) found female talkers were more intelligible, so these differences in duration, perhaps due to articulatory flexibility, could potentially impact the intelligibility of speech.

Results for vowel dispersion data did not show the expected pattern, and gender had no effect in either by-speaker or by-vowel dispersion measures. Although not statistically significant, a marginal effect in the by-vowel vowel dispersion measure indicates vowels for low frequency words were produced with more variation than vowels in high frequency words. The direction of this effect may indicate that low frequency words result in larger by-vowel dispersion, perhaps as a consequence of increased effort or because the lexical items are unfamiliar and have fewer associations with habitual pronunciations.

Other stimulus characteristics may have affected dispersion results, as our results differed from those of previous studies that found gender and word frequency effects differences on

vowel dispersion measures (Bradlow et al., 1996; Pardo, Urmanche, Wilman, & Wiener, 2017b). While the present study used bisyllabic words, previous research finding significant effects of gender and of lexical category in vowel dispersion has used a variety of other kinds of stimulus materials. For example, Bradlow et al. (1996) found females had higher vowel dispersion using sentences-length utterances. Although it is possible that bisyllabic words had insufficient variation to be measured, it seems unlikely considering that other studies have found differences in vowel dispersion for monosyllabic words. Munson and Solomon (2004) found low frequency monosyllabic words had higher dispersion, and Pardo et al. (2013) replicated those findings. Pardo et al. (2017b) found similar results using both bisyllabic and monosyllabic word stimuli (from Munson & Solomon, 2004). An alternative possibility for my study's results is the specific vowel content of the stimulus items.

One of the main goals of the present study was to assess the role of stimulus properties in driving differences in pronunciation between males and females. The vowel categories explored by the present study, /a, ɛ, u, i/, differed from the vowels examined by previous studies, and could be a reason for the diverging results. Bradlow et al. (1996) only used the vowels /a, i, o/, and purposefully excluded the vowel /u/ because it possesses multiple forms or variations in the English language. Munson and Solomon's analysis (2004) included the same vowels as in this study, as well as the vowels /æ, eɪ, oʊ, ʊ/. Thus, differences in stimulus properties, including the vowel's qualities, might have affected the results. Particularly in the dispersion analysis by-speaker, the present study may have reduced the sources of variation by including four point vowels or increased within-category variation by including sounds that are too dispersed in the English language. It could also be the case that low power may have reduced the ability to detect true effects. Although the gender differences in pitch and the gender by word frequency

interaction on vowel duration were robust, the effects of gender and word frequency on vowel dispersion may be subtler.

The Lobanov technique was used for vowel normalization, and upon more careful review, this technique might have removed some of the desired variation from the utterances. Although previous research has used this technique (Babel, 2012) and succeeded in preserving vocal imitation information, studies comparing normalization techniques (e.g., Fabricius et al. (2009) found the Lobanov technique may remove some of the variation attributed to speaker sex. Since the present study aimed at analyzing any existing sex differences between males and females in several measures, the effects might have been lessened by removing some of the gender variation when normalizing vowels. Furthermore, the Lobanov technique is considered *speaker*-intrinsic, meaning it normalizes vowel data from one speaker at a time. The Labov technique (Labov, Ash, & Boberg, 2006), by contrast, uses speaker information from a population to normalize vowels for every speaker, which might have been more appropriate when conducting our analysis (Fabricius et al., 2009). Future investigations of the phenomena examined in this study should consider using other normalization techniques, or calculating vowel dispersion measures on formant values that have not been normalized.

In order to investigate whether the gender and word frequency effects in vowel and word duration are in fact due to articulation flexibility, future research should include a perceptual analysis of the utterances used here. Bradlow et al. (1996) showed speakers' vowel dispersion affected the intelligibility of their utterances, and those with higher dispersion were found to be more intelligible. The hypothesis, following Bradlow et al. (1996), is that talkers who vary their pronunciations as a function of lexical characteristics such as word frequency will be more intelligible. Therefore, if the word frequency effect in vowel and word duration is in fact related

to articulation flexibility, then the speakers and utterances with longer durations should be the most intelligible.

Future research should also consider expanding on the idea that these differences in pronunciation are meant to *perform* gender. If any differences in males' and females' pronunciation as reflected in acoustic-phonetic properties such as duration and dispersion influence the perception of speech and of the talker, one could make the case that speakers have been successful in performing gender. Additionally, by collecting ratings of voices on their gender stereotypicality (Babel & McGuire, 2015; Johnson, Strand, & D'Imperio, 1999), I could investigate whether duration has an effect on the speaker's perceived gender, such that speakers with faster speaking rates are perceived as more masculine, for example.

In conclusion, the study of speech production can expand our knowledge of the interplay between social characteristics like gender and linguistic characteristics such as word frequency. While dispersion data had no significant results according to speaker gender, there were significant effects of both word frequency and gender on duration measures, suggesting that males and females may differ in their production of spoken language. In order to be fully understood in the context of social communication, these findings could also be investigated from a speech perception and intelligibility perspective. The current findings reveal an important difference between males and females in their sensitivity to the properties of words. The present study informs our understanding of how the voice can be used to perform gender, why gender differences may exist in vocal accommodation and convergence, and the role of gender in language change over time.

References:

Adank, P., Smits, R., & Hout, R. v. (2004). A Comparison of Vowel Normalization Procedures for Language Variation Research. *Journal of the Acoustical Society of America, 116*(5), 3099-3107. doi:10.1121/1.1795335

Babel, M. (2012). Evidence for Phonetic and Social Selectivity in Spontaneous Phonetic Imitation. *Journal of Phonetics, 40*(1), 177-189. doi:10.1016/j.wocn.2011.09.001

Babel, M., & Bulatov, D. (2012). The Role of Fundamental Frequency in Phonetic Accommodation. *Language and Speech, 55*(2), 231-248. doi:10.1177/0023830911417695

Babel, M., & McGuire, G. (2015). Perceptual Fluency and Judgments of Vocal Aesthetics and Stereotypicality. *Cognitive Science, 39*(4), 766-787. doi:10.1111/cogs.12179

Balota, D., Yap, M., Hutchison, K., Cortese, M., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*(3), 445-459. doi:10.3758/BF03193014

Bilous, F. R., & Krauss, R. M. (1988). Dominance and accommodation in the conversational behaviours of same- and mixed- gender dyads. *Language and Communication, 8*(3), 183-194. doi:10.1016/0271-5309(88)90016-X

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer (Version 6.0.36). http://www.praat.org/.

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine- grained acoustic- phonetic talker characteristics. *Speech Communication, 20*(3), 255-272. doi:10.1016/S0167-6393(96)00063-5

Burgess, C., & Livesay, K. (1998). The effect of corpus size in predicting reaction time in a basic word recognition task: Moving on from Kučera and Francis. *Behavior Research Methods, Instruments, & Computers, 30*(2), 272-277. doi:10.3758/BF03200655

Chiffi, D. (2012). Idiolects and Language. *Axiomathes, 22*(4), 417-432. doi:10.1007/s10516-011-9151-2

Diehl, R. L., Lindblom, B., Hoemeke, K. A., & Fahey, R. P. (1996). On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics, 24*(2), 187-208. doi:https://doi.org/10.1006/jpho.1996.0011

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language variation and change, 1*(3), 245-267.

Eckert, P., & McConnell-Ginet, S. (1992). Think Practically and Look Locally: Language and Gender as Community- Based Practice. *21*, 461-490.

Fabricius, A. H., Watt, D., & Johnson, D. E. (2009). A Comparison of Three Speaker- Intrinsic Vowel Formant Frequency Normalization Algorithms for Sociophonetics. *Language variation and change, 21*(3), 413-435. doi:10.1017/S0954394509990160

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America, 106*(3), 1511-1522. doi:10.1121/1.427148

Flynn, N., & Foulkes, P. (2011). *Comparing vowel formant normalization methods.* Paper presented at the Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII).

Fraccaro, P. J., amp, Amp, Apos, Connor, J. J. M., Re, D. E., . . . Feinberg, D. R. (2013). Faking it: deliberately altered voice pitch and vocal attractiveness. *Animal Behaviour, 85*(1), 127-136. doi:10.1016/j.anbehav.2012.10.016

Haas, A. (1979). Male and female spoken language differences: Stereotypes and evidence. *Psychological Bulletin, 86*(3), 616-626. doi:10.1037/0033-2909.86.3.616

Inc., F. T. (2017). Sound Studio (Version 4.8.13). www.felttip.com.

Johnson, K. (2005). Speaker normalization in speech perception. *The handbook of speech perception*, 363-389.

Johnson, K. (2006). Resonance in an Exemplar- Based Lexicon: The Emergence of Social Identity and Phonology. *Journal of Phonetics, 34*(4), 485-499. doi:10.1016/j.wocn.2005.08.004

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory- Visual Integration of Talker Gender in Vowel Perception. *Journal of Phonetics, 27*(4), 359-384. doi:10.1006/jpho.1999.0100

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*(5), 580-602. doi:10.1016/0749-596X(91)90027-H

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating Variation: Dialects, Idiolects, and Speech Processing. *Cognition, 107*(1), 54-81.

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English : phonetics, phonology, and sound change : a multimedia reference tool*. Berlin ; New York: Berlin ; New York : Mouton de Gruyter.

Lakoff, R. T. (2004). *Language and Woman's Place Text and Commentaries* (2nd ed.. ed.). Oxford: Oxford : Oxford University Press, USA.

Lewandowski, E. (2018). *The Automaticity of Entrainment: Perceptual, social and individual differences in vocal alignment.* (PhD Dissertation), Emory University.

Li, T., & Fu, Q.-J. (2011). Perceptual Adaptation of Voice Gender Discrimination with Spectrally Shifted Vowels. *Journal of Speech, Language, and Hearing Research, 54*(4), 1240-1245. doi:10.1044/1092-4388(2010/10-0168)

Lobanov, B. M. (1971). Classification of Russian Vowels Spoken by Different Speakers. *The Journal of the Acoustical Society of America, 49*(2B), 606-608. doi:10.1121/1.1912396

Lund, K., & Burgess, C. (1996). Producing high- dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*(2), 203-208. doi:10.3758/BF03204766

Macaulay, R. (2010). *Dialect*. Amsterdam, Netherlands: Amsterdam, Netherlands: Benjamins.

Miller, R. M. (2011). Talker- Specific Accent: The Role of Idiolect in the Perception of Accented Speech.

Mullennix, J. W., Johnson, K. A., Topcu-Durgun, M., & Farnsworth, L. M. (1995). The Perceptual Representation of Voice Gender. *Journal of the Acoustical Society of America, 98*(6), 3080-3095. doi:10.1121/1.413832

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of speech, language, and hearing research : JSLHR, 47*(5), 1048. doi:10.1044/1092-4388(2004/078)

Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender Differences in Vocal Accommodation. *Journal of Language and Social Psychology, 21*(4), 422-432. doi:10.1177/026192702237958

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes, 45*(3), 211-236.

Nygaard, L. C., & Queen, J. S. (2000). The role of sentential prosody in learning voices. *The Journal of the Acoustical Society of America, 107*(5), 2856-2856.

Owren, M. J., & Bachorowski, J.-A. (2003). Reconsidering the Evolution of Nonlinguistic Communication: The Case of Laughter. *Journal of Nonverbal Behavior, 27*(3), 183-200. doi:10.1023/A:1025394015198

Pardo, J., Jay, I., & Krauss, R. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics, 72*(8), 2254-2264. doi:10.3758/BF03196699

Pardo, J., Jordan, K., Mallari, R., Scanlon, C., & Lewandowski, E. (2013). Phonetic Convergence in Shadowed Speech: The Relation between Acoustic and Perceptual Measures. *Journal of Memory and Language, 69*(3), 183-195. doi:10.1016/j.jml.2013.06.002

Pardo, J., Urmanche, A., Wilman, S., & Wiener, J. (2017a). Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics, 79*(2), 637-659. doi:10.3758/s13414-016-1226-0

Pardo, J., Urmanche, A., Wilman, S., & Wiener, J. (2017b). Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics, 79*(2), 637-659. doi:10.3758/s13414-016-1226-0

Perkell, J. S., Zandipour, M., Matthies, M. L., & Lane, H. (2002). Economy of Effort in Different Speaking Conditions, I: A Preliminary Study of Intersubject Differences and Modeling Issues. *Journal of the Acoustical Society of America, 112*(4), 1627-1641. doi:10.1121/1.1506369

Podesva, R. J., & Hofwegen, J. V. (2015). */ s/ exuality in Smalltown California: Gender Normativity and the Acoustic Realization of / s*: Oxford University Press.

Pollack, I., Pickett, J. M., & Sumby, W. H. (1954). On the Identification of Speakers by Voice. *The Journal of the Acoustical Society of America, 26*(3), 403-406. doi:10.1121/1.1907349

Provine, R. R. (1993). Laughter punctuates speech: Linguistic, social and gender contexts of laughter. *Ethology, 95*(4), 291-298.

Provine, R. R. (2017). Laughter as an approach to vocal evolution: The bipedal theory. *Psychonomic Bulletin & Review, 24*(1), 238-244. doi:10.3758/s13423-016-1089-3

RCoreTeam. (2016). R: A Language and Environment for Statistical Computing (Version R 3.3.1 GUI 1.68 Mavericks build). https://www.R-project.org: R Foundation for Statistical Computing.

Rendall, D., Vasey, P., & McKenzie, J. (2008). The Queen's English: An Alternative, Biosocial Hypothesis for the Distinctive Features of " Gay Speech". *Archives of Sexual Behavior, 37*(1), 188-204. doi:10.1007/s10508-007-9269-x

Shin, N. L. (2013). *Women as leaders of language change: A qualification from the bilingual perspective.* Paper presented at the Proceedings of the 6th International Workshop on Spanish Sociolinguistics.

Sneller, B., & Roberts, G. (2018). Why some behaviors spread while others don't: A laboratory simulation of dialect contact. *Cognition., 170*, 298. doi:10.1016/j.cognition.2017.10.014

Strand, E. A., & Johnson, K. (1996). *Gradient and visual speaker normalization in the perception of fricatives.* Paper presented at the KONVENS.

Stuart-Smith, J. (2007). *Empirical Evidence for Gendered Speech Production: / S/ in Glaswegian.* Paper presented at the Laboratory Phonology, Berlin, Germany.

Sweller, J., & Jose, P. M. a. B. H. R. (2011). CHAPTER TWO - Cognitive Load Theory *Psychology of Learning and Motivation* (Vol. Volume 55, pp. 37-76): Academic Press.

Titze, I. R. (1989). Physiologic and Acoustic Differences between Male and Female Voices. *Journal of the Acoustical Society of America, 85*(4), 1699-1707. doi:10.1121/1.397959

Appendix: Stimulus Words and Word Frequency

High frequency words				Low frequency words			
Word	HAL Frequency	Word	HAL Frequency	Word	HAL Frequency	Word	HAL Frequency
anon	18879	monster	15607	abscond	17	meager	771
apart	22038	motif	15926	arson	752	menace	1308
assume	42861	movement	30554	banal	436	misdeal	12
cannot	108318	movie	61248	bardic	99	miso	213
center	87347	neither	37284	bazaar	2370	moonlit	184
connect	37477	never	303028	beady	82	narwhal	218
dealer	17655	people	768168	benzene	235	nocturne	122
demo	24690	placement	10692	burble	57	octet	169
disease	23168	percent	25919	cadet	1017	offend	2914
duty	14997	reduce	21540	cahoots	161	peaky	45
event	42825	regard	15551	chenille	64	poodle	743
female	43053	remove	43954	demean	307	preheat	509
follow	58490	resume	80442	denude	12	raccoon	275
forget	43510	review	39843	divest	88	seduce	545
garbage	12598	robot	10880	draftee	39	senile	543
indeed	43972	special	105765	ensconce	8	sensual	1455
into	563310	student	56461	foment	125	snooty	142
instead	106318	super	28262	gavotte	53	stocking	1433
involve	10020	topic	34583	ghoulish	49	supine	170
issue	117620	unique	24825	harlot	164	vamoose	10
machine	121109	vendor	12408	imam	1074	vegan	1663
marvel	19626	volume	40289	immense	2256	vengeance	1999
meeting	37663	workshop	15377	kazoo	148	volley	548
member	63948	zero	21278	meadow	841	voodoo	3326