Voice Characteristics of Transgender Speakers

By Rachael Harrington

B.A. in French Language and Literature, May 2009, The George Washington University

A Thesis submitted to

The Faculty of
the Columbian College of Arts and Sciences
of The George Washington University
in partial fulfillment of the requirements
for the degree of Master of Arts

May 15, 2011

Thesis directed by

Adrienne Hancock
Assistant Professor of Speech and Hearing Sciences

Acknowledgments

My thanks go first and foremost to Dr. Adrienne Hancock for her calming force in guiding me through the research and writing process. I could not have finished, interpreted, or written this project without you. You pushed me outside of my comfort zone and made me think about and analyze the project in new and exciting ways.

Thank you to my committee, Linda Siegfriedt and James Mahshie, for all of your insightful comments and critiques. Thank you to Professor Siegfriedt for introducing me to voice therapy and this population. Thank you to Dr. Mahshie for teaching me the techniques I needed to navigate the complicated world of acoustic and aerodynamic analysis.

Thank you to the members of the voice lab: Amy Lopez for your design, Kelsey Day, Laura Garabedian, and Erica Altman for all of your help with data collection and analysis, and Laura Sibbald for being a great reader and friend.

A million thanks go to the undergraduate and graduate students and faculty of The George Washington University Speech and Hearing Sciences Department for their great spirits, compassion, laughs, and support.

I would also like to thank my parents and non-SPHR friends. Thank you to my mom and dad, Mary and Chip, for going above and beyond to help me complete this project with constant love, insightful reading, and travel. I could not have finished without you. Thank you to all of my friends for reading and forwarding my many emails and helping to recruit volunteers. Thank you to Jake, for spending endless hours keeping me company in the lab and just being generally wonderful.

Abstract of Thesis

Voice Characteristics of Transgender Speakers

Transgender voice therapy is often designed by considering cisgender male and female gender differences and some limited research about effects of voice measures on gender perception. Vocal fold vibration patterns are known to differ between genders but are not traditionally included as targets in voice feminization. This study aims to describe and explain these vocal fold vibration patterns to determine if the transgender larynx modifies its patterns to be more like those of the desired gender. Data was collected from ten cisgender males, ten cisgender females, and ten male-to-female transgender speakers. Derived subglottal pressure, airflow amplitude (AC flow, peak flow, and DC flow), airflow duration (open quotient, closed quotient, open speed quotient, aQuotient, and bQuotient), and glottal closure duration (closed speed quotient, cQuotient, and dQuotient) values were calculated from this data. It was hypothesized that airflow amplitude values would differ significantly from either gender and that duration values would approximate the values of the desired gender. Statistically significant differences were found in the values of peak flow (between males and females), AC flow (between males and females), open and closed quotients (between males and females and males and male-to-female transgender speakers), and aQuotient (between males and male-to-female transgender speakers). The data suggest that male-to-female transgender speakers' vocal folds are open for a longer percentage of the total duration of the cycle than either male or female cisgender speakers' vocal folds; this appears to be achieved by increasing the length of time during which the vocal folds are opening. It can be concluded from the data that this increase is achieved with a unique vocal fold closure pattern that is different from, and

not in between or approximated to, either male or female cisgender vocal fold closure patterns.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction and Literature Review

This study aims to examine the glottal source characteristics of transgender (TG) speakers in order to understand the phenomenon of when a biologically male larynx is the source of a female-sounding voice. There are three possible scenarios for how the functioning of the larynx of an MTF person changes to create a voice more like that produced by biological females. One is that they modify their airflow or glottal configuration to approximate a value *between* male and female values, perhaps unable to achieve female values but getting close enough for perception as female. This explanation is similar to the gender-neutral range supported by the literature for F_0. A second explanation is that TG speakers increase or decrease airflow or glottal configuration measure values to the extent that they do *approximate the desired gender's* airflow and duration measure values. That is, the values of MTF speakers are able to approximate female values and would only be significantly different from male values. If the MTF speaker achieves values at the desired gender's values or between the cisgender values for one category of measures (i.e. airflow amplitude), then they will likely require a third explanation for the pattern of the other category (duration). The third explanation is that MTF speakers achieve values that are *completely different* than either males or females because they are using unique compensation strategies to create a female voice with a male larynx.

Speaking fundamental frequency and formant measures of TG speakers typically progress from male-like toward female-like as they become recognized as sounding female. Airflow data from one study suggest that MTF airflow will equal or will exceed the published normative data for female speakers rather than stay between the male and female values (Gorham-Rowan & Morris, 2006). It is possible that the MTF speaker achieves a female-sounding voice using a laryngeal physiology distinct from cisgender (i.e., non-TG) females to create these increased airflow values. Glottal source characteristics are a relatively unexplored area of TG research, but it should not be assumed that airflow and vocal fold configuration measures of TG voice are like F0 and formants (i.e., simply between those of male and female voice).

The purpose of this study is to report the derived subglottal pressure, airflow amplitude, and airflow and electroglottogram timing measure values of MTF speakers. To understand the following hypotheses for MTF voice it is necessary to review the physiology of vocal fold vibration, measures of glottal source characteristics, and published values for glottal source characteristics of cisgender and TG voice.

Transgender Communication

The American Psychological Association (APA, 2000) has two criteria for a classification of gender identity disorder. The first criterion is “evidence of a strong and persistent cross-gender identification.” The second criterion is “evidence of persistent discomfort about one's assigned sex.” Transgender is a population classified within this diagnosis. Adler, Hirsch, and Mordaunt (2006) define the term transgender as an “umbrella definition used for an individual who is living in a gender other than his or her own biological gender” (p. 5). These individuals strive to look, sound, and be perceived as the desired gender. Typically, TG individuals will take hormones to better achieve these goals. For female-to-male (FTM) speakers, androgen hormones (usually testosterone) increase the size of the larynx and thicken the muscles of the vocal folds, leading to a deeper, more masculine voice. This is not the case, however, for all types of hormone therapy. Andrews (2006) states that male-to-female (MTF) hormones, such as estrogen and progesterone, do not affect the size or makeup of the larynx and, as a result, MTF individuals typically seek speech-language pathology (SLP) services to achieve a more feminine voice. A more feminine voice is achieved by altering several aspects of TG communication to approximate female values, at least to some extent.

The SLP scope of practice (American Speech, Language, and Hearing Association [ASHA], 2007) includes the feminization of TG voice as a subset of communication enhancement or modification. Speech-language pathologists are well-qualified for the feminization of TG voice as the process for TG voice feminization uses many of the same techniques that are used in the treatment of traditional voice populations (e.g., forward focus and intonation therapy). Voice feminization without therapy can be a vocally abusive behavior, so an SLP may ultimately treat a voice feminized without therapy because this puts speakers at risk for voice disorders (ASHA, 2011). Several authors and organizations have recommended increasing the amount of available research on techniques currently used in TG voice therapy while others have used the currently available literature to recommend treatment plans for the healthy feminization of TG voice.

ASHA (2011) lists the following target areas for approximation or achievement of female values and characteristics in voice feminization therapy: “pitch, resonance, intonation, rate, volume/intensity, language, speech sound production (articulation) and pragmatics (social rules of communication).” Gelfer (1999) states that therapy should terminate only when the TG client and clinician are comfortable with “pitch, quality, intensity, variability, and natural-sounding intonation patterns in spontaneous speech” (p.

206). Adler et al. (2006) list the following target areas: “pitch, resonance, intonation, rate, volume, syntax, vocabulary, pragmatics, articulation, vocal hygiene, nonverbal communication, carryover, singing voice” (p. 13). Gelfer specifically focuses on aspects of traditional voice therapy while Adler et. al. and ASHA include verbal and nonverbal language and articulation as well.

Oates and Dacakis (1983) wrote the first overview article about TG voice therapy. At the time, only case reports of TG speakers increasing speaking fundamental frequency suggested that the TG voice could be modified. Oates and Dacakis recommended research in anatomical and physiologic constraints, speakers’ goals and clinicians’ biases, and how nonverbal factors, psychosocial factors, and diagnostic status may influence therapy. To date, much treatment research is limited to case reports and so relies heavily on research about what voice characteristics contribute to gender perception, particularly pitch and resonance. For those voice characteristics, it is hypothesized that a MTF speaker is more likely to be identified as female as the values become more like those of cisgender female speakers (e.g., perceived femininity increases as pitch increases from male range toward female range). Feminizing communication by approximating or achieving female normative values on a variety of communication measures is logical and is supported by clinical case reports in the literature. Some research is available to support these targets based on correlations with listener perception of gender. Ultimately, voice feminization treatment effectiveness is measured by listeners’ perceptions. Some authors suggest approximating female values with a gender-neutral fundamental frequency (F_0), or a range in which gender is inconsistently attributed from perception of pitch, and achieving formant frequencies at or above female values for the most effective, healthy voice feminization. It is presumed that when a TG speaker has a gender-neutral pitch, hearers will assign gender on the basis of other perceptual cues.

Pitch is the only aspect of TG voice treatment with evidence in the literature to suggest a gender-neutral range is effective for perception as the desired gender. Wolfe, Ratusnik, Smith, and Northrop (1990) report the average F_0 for an FTM speaker perceived as male is 118 Hz and an FTM speaker perceived as female is 171 Hz, somewhat above and below cisgender normative values. Gelfer and Shofield (2000) report the average F_0 for an MTF speaker is 187 Hz with a range of 164 to 199 Hz, or on the low end or slightly below cisgender female values. The F_0 required to be perceived as gender-neutral, however, is in a significantly smaller range. The lowest F_0 of speakers perceived as gender-neutral is reported by Wolfe et al. as 155 Hz and by Spencer (1988) as 160 Hz. In a compromise, Adler et al. (2006) report the most gender-neutral F_0 to be between 155-160 Hz. Targeting pitch in this gender-neutral range for MTF speakers, in combination with other techniques like forward-focus to change resonant voice quality, has proven to be successful in creating a perceptually feminine voice (Andrews & Schmidt, 1997; Holmberg, Oates, Dacakis & Grant, 2009; Childers & Wu, 1991; Gelfer & Mikos, 2004; Gelfer & Schofield, 2000). While approximating female values with a gender-neutral pitch is more feasible than achieving the values for pitch of a cisgender female that is not the case for other areas of communication. Some areas, like formant frequencies, do require that TG speakers achieve values more similar to female levels to be effectively perceived as female.

Formant frequencies are resonances created within the vocal tract that are modulations of the initialF_0 created at the level of the larynx. The three formants are referred to as F_1, F_2, and F_3. F_1 modulates frequency based on the relative opening of the oral cavity. F_2 modulates frequency based on the relative anterior-posterior placement of the tongue. F_3 modulates frequency based on the relative length of the oral cavity (Kent, 1997). Hillenbrand and Clark (2009) found that more male speakers are identified as feminine when all formants and F_0 are digitally shifted from produced values to achieve normative female values, compared to only shifting F0 or only shifting formants. Mount and Salmon (1988), report that a MTF speaker was not identified as female until they achieved formant frequencies consistent with feminine normative data for formant frequencies, even though F_0 was within the gender-neutral range or within female normative values. Gelfer and Schofield (2000) found that achieving formant frequencies similar to female normative data correlates to perception of speaker femininity. These results should be interpreted with caution, however, because results are based on 3 out of 13 MTF speakers who were perceived as female and the results were not statistically significant. Overall, several sources suggest that TG speakers should achieve formant frequencies similar to those of a female, as opposed to approximating the values, for a voice to be perceived as female. It is unclear if any particular formant contributes to this perception more than any other (Childers & Wu, 1991; Gelfer & Mikos, 2005). All of the pitch and formant frequency data support a commonly used assumption, to be referred to as the typical TG hypothesis, that the MTF speaker needs to approach or meet the female normative values in order to be perceived as a female speaker.

Gender Differences in Voice

Apart from pitch and resonance, most rationale for TG voice feminization is based on known gender differences. Some of these gender differences arise because the male and female larynxes are anatomically and functionally different. Boone (2010) states that the male larynx sits low in the neck, which causes the resonances of the vocal tract to be generally lower in frequency compared to those of females. Further, male vocal folds are longer and thicker than those of females', which causes the F_0 of the voice to be lower in pitch. Male speaking F_0 is reported on average to be 120 +/-20 Hz, whereas female is reported on average to be 220 +/- 20 Hz (Andrews, 2006). The tension of the vocal folds also contributes to F_0.

The size, shape, and placement of the larynx and vocal folds also affect loudness. Male voices tend to have higher intensity levels than females, however, average conversational levels for both genders range between 50-65 dB (Adler et. al, 2006). More specifically, Oates and Dacakis (1983) report that female average intensity level is 64 dB and male average intensity level is 67 dB. Increased subglottal pressure, or increased pressure below the vocal folds, is the main variable in increasing intensity; however it can be used to alter other aspects of voice. The male vocal folds are larger, which can grant a greater range of control over subglottal pressure, and this may account for the gender differences. The comparative difference in size of the vocal folds and vocal tracts of the two genders also leads to a difference in formant frequencies across vowels.

There are gender differences specific to formant frequencies. Fundamental frequency, F_1, F_2, and F_3 all have greater values, perhaps as much as 20% greater, for females than for males, because of the smaller vocal tract size of females (Coleman, 1983). Formants can be slightly increased by elongating the vocal tract with a relatively more open and forward tongue positioning, which is a common treatment goal in voice feminization. Kent (1997) combines data from Peterson and Barney (1957) and

Hillenbrand, Getty, Clark, and Wheeler (1995) to show gender differences in all formant frequencies during all vowels. Gunzberger (2004), however, states that F_3 is typically the only formant that is consistently different across vowel types between the two genders. The configuration of the vocal fold differs between genders, like the configuration of the vocal tract.

The configuration of vocal fold closure is another difference between male and female speakers. According to Sodersten and Lindestad (1990), women, specifically women aged 20-35 yrs, may occasionally show a posterior glottal chink, or an opening in the posterior third of the larynx, during phonation. Hixon, Weismer, and Hoit (2008) state that men, however, will typically achieve complete closure across the entire larynx. Linville (1992) mentions the possibility that this posterior glottal chink is an unintended result of an intentional behavior to create a perceptually breathy voice. A breathy voice is perceptually associated with a feminine voice (Linville, 1992; Klatt & Klatt, 1990).

With possible exception of F_0 and formant frequencies, it is unclear if it is appropriate to establish targets in the areas recommended for voice and communication feminization by ASHA (2011) based on gender specific normative data. The typical TG hypothesis suggests that those targets should approximate or achieve values of the desired gender so that a TG speaker is more likely to be perceived as the desired gender. Preliminary studies into the glottal characteristics of MTF speakers, however, suggest that this may not be the case. Gorham-Rowan and Morris (2006) indicate that there are significant differences in glottal source characteristics for TG and cisgender speakers in which the TG speaker values do not approximate or achieve the desired gender values, but create values entirely different from either gender. Collection of more data on glottal source characteristics is a first step to determining if there is a reason to consider adding glottal vibration cycles to the list of therapy targets for TG voice feminization and whether these characteristics should be targeted in the vein of the typical TG hypothesis (achievement or approximation) or in an entirely different manner as is suggested by Gorham-Rowan and Morris.

Vocal Fold Vibration

The source-filter theory describes how sound is produced and modulated within the vocal tract (Ferrand, 2007; Kent, 1997). Created by Gunnar Fant in the 1960s, the source-filter theory explains how air from the lungs becomes sound. The source-filter theory has three parts: the source function, the transfer function, and the output function. For vowels, the source function refers to vocal fold movement patterns used to create sound. The vocal folds trap air in the lungs and then release that air to create the F_0 and spectrum of a vowel sound (Ferrand). As the vocal folds are held together with surface tension and muscular force, pressure builds below them in the lungs. When the pressure below the vocal folds (i.e., subglottal pressure) exceeds the pressure above the vocal folds, the vocal folds open and, during voiced sounds, vibrate as the air passes them. If, for example, the vocal folds are being held together with a large amount of muscular force, then the subglottal pressure will increase to be able to exceed that force and open the vocal folds. Subglottal pressure is typically manipulated to increase loudness; however, it can also make adjustments in response to vocal fold tension and airflow rate. The transfer and output functions modulate this air into the sound that is heard as speech.

Characteristics of Vocal Fold Vibration

The vocal folds vibrate in a specific pattern. As the vocal folds open, a mucosal traveling wave begins within the layers. The inferior and superior edge of the vocal folds, while both influenced by the same mucosal wave, open at different times and in different manners, creating a vertical phase difference. The inferior edge begins to open prior to the superior edge and reaches its maximum excursion while the superior edge is still opening. The superior edge opens after the inferior edge and reaches its maximum excursion as the inferior edge has already begun to close. Both edges open anteriorly to posteriorly in a cycle (Kent, 1997). This pattern can make it hard to define the open and closed phases of the cycle. To visualize these parts of the cycle, it is best to derive a glottogram, or a visual representation of the cycles (*see* Figures 1 and 2). A glottogram can be derived from pneumotachography or electroglottography (EGG).

The airflow glottogram, in the top portion of Figure 1, is a representation of the amount of airflow, measured in L/s, over time, measured in seconds. The peaks and valleys of the glottogram are created by obstruction to air flowing from the lungs created by vocal fold vibration. Airflow decreases (valley of airflow glottogram) as the vocal folds close and narrow the glottis; it increases (peak of the airflow glottogram) as the vocal folds open the glottis. The electroglottogram in the bottom portion of Figure 1 is a representation of the amount of conductance over time, measured in seconds by an electroglottograph. An electrical current travels more easily through tissue than it does through air, so areas of highest conductance, or when the vocal folds are the most approximated, are represented as peaks, and areas of lowest conductance, or when the vocal folds are least approximated, are represented as valleys.

No matter how a glottogram is derived, it can be divided into four quarter segments (*see* Figure 1). First, the waveform is divided into two phases, open and closed, at the point on the x-axis between the peak and the valley when the waveform crosses a criterion line. The criterion line used for this particular study is drawn to equal 15% of the total amplitude of the cycle above the lowest point in the cycle on the y-axis (T_2 in Figure 1; Rothenberg & Mahshie, 1988). The 15% criterion line was originally chosen because it includes an adequate amount of the area around the peak without confounding it with other factors (neck thickness or electrode placement for EGG or vocal fold closure for airflow), however, other criterion lines are regularly employed and are employed in studies referenced later (e.g. Holmberg, Hillman, Perkell and Gress (1994) which employs a 30% criterion line as was suggested by Sapienza, Stathopoulous, and Dromey (1998)). The 15% criterion line is best for this particular study because it separates reasonably well the opened and closed portions of the glottal vibratory cycle and is used by many of the articles used to design hypotheses (Higgins & Saxman, 1991; Chen et. al., 2002). Using a similar method of determining open and closed phase is useful in comparing data across studies. This criterion line also determines the start and end (T_0 in Figure 1) of each cycle. On the airflow glottogram, the waveform above this line represents when the vocal folds are open. On the EGG, the waveform above this line represents when the vocal folds are closed.

These two halves of the cycle are further divided into segments at the point of the peak (T_1 for airflow and T_3 for EGG in Figure 1) and valley (T_3 for airflow and T_1 for EGG in Figure 1) of each cycle. Figure 1 displays these four segments as a, b, c, and d. The aSegment represents a period of time from when the vocal folds are opening to the time when they reach the most open point in the cycle. The bSegment represents a period of time from when the vocal folds are most open to the time when they are considered closed. The cSegment represents time from the beginning of the closed segment, when the vocal folds are closing, to the point in the cycle when they are the most closed. The dSegment represents a period of time from when the vocal folds are the most closed to when they are considered open.

An airflow glottogram gives detailed information about amplitude and temporal values in the open phase of each vibration cycle (a+b) but no useful information about the closed phase (c+d) of the vocal fold vibration cycle because airflow is occurring minimally when the vocal folds are closed. This information about the closed phase is not useful because the only airflow that is being recorded is airflow that is leaking from closed vocal folds. The EGG, when gathered simultaneously with airflow during speech production, can accurately depict the temporal measure of vocal fold movement during these c and d segments; EGG tells little about the open portion of the vocal fold vibration cycle because change in conductance is uninformative when the vocal folds are open. Airflow and EGG glottograms of the same cycle at the same point in time supplement one another and create a complete representation of the glottal cycle.

Measurement of Glottal Source Characteristics

Temporal and amplitude values are gathered from the airflow glottogram. Amplitude measures, expressed in L/s, include peak flow, DC flow and AC flow. Peak flow is the amplitude of the signal at T_1, less any drift artifact. Perkell, Hillman, Holmberg, and Gress (1994) report that peak flow is higher for males than for females. DC flow is the difference between the amplitude value of T_3 on Figure 1 and zero, representing the constant, but turbulent, airflow that is not modulated by the vocal folds opening and closing to create pulses of airflow. Perkell et al. and Higgins and Saxman (1991) found that DC flow was greater for females than for males; however, the differences for both studies were small and insignificant. AC flow is the difference between peak flow and DC flow. Table 1 describes these measures, their calculations, and normative values in greater detail.

The airflow signal is also measured temporally. The open quotient is the duration of time when the vocal folds are open (a + b in Figure 1) divided by the total period of the cycle. Conversely, the closed quotient is the duration of time when the vocal folds are closed (c + d in Figure 1), divided by the total period. Higgins and Saxman (1991) report that the open quotient is larger in females than in males and that the closed quotient is larger in males than in females, indicating that the female vocal folds are open for a longer percentage of time during phonation as compared to males. Both quotients are calculated from the airflow glottogram, despite the fact that the airflow glottogram does not give the most complete information about the closed phase, because the closed quotient from the EGG glottogram may not compare exactly to the open quotient from an airflow glottogram. This is because a different criterion line is calculated for airflow and EGG. Theoretically, they should be the same as they are calculated from the same points in time with the same method; however, there are often slight differences because the amplitude of the EGG signal is dependent on factors outside of glottal closure, such as neck thickness.

As described above, each phase of the cycle is divided into quarter quotients (a, b, c, and d in Figure 1). The phase of the cycle above the criterion line (15% in this study) on the airflow glottogram is divided into the aSegment and the bSegment. The aQuotient is the percentage of the aSegment of the cycle's period. No normative data are available for this quotient. The bQuotient is the percentage of the bSegment of the cycle's period. Sulter and Wit (1996) report this value to be greater in females than in males. The aSegment is divided by the bSegment to calculate the open speed quotient (OSQ). A symmetrical cycle will have an OSQ value close to one. A low speed quotient value is created by an opening phase much faster than the closing phase and this may indicate tension in the vocal folds. Sulter and Wit report that this value is greater than one for both genders, but greater in males than females. This indicates that more time is spent opening than closing the vocal folds during the open segment for both genders, more so for males. For definitions of time measures in an airflow wave and previously reported numerical values for males and females in greater detail please refer to Table 2.

Electroglottograms are measured for time only because amplitude cannot be calibrated. Electroglottograms provide information about the portion of vocal fold vibration when the vocal folds are approximated. The cQuotient is the duration of time of cSegment divided by the period. Chen, Robb, and Gilbert (2000) report the cSegment in ms and do not convert it into a quotient. All data regarding these small quarters should be reported with quotient because it is necessary to account for variable period durations. This is especially important for comparing across genders because they have significantly different F_0 . Chen et. al reported cSegment to be greater for men than for women, and when corrected for quotient by calculating with an average F_0 reported for males and females, it can be derived that they find cQuotient to be greater for men than for women. The dQuotient is the period of time of dSegment divided by the period. Chen et al. report dSegment to be greater in women than in men. When corrected for quotient by calculating with an average F_0 reported for males and females, it can be derived that they find dQuotient to be greater for women than for men. The dSegment is divided by the cSegment to determine the closed speed quotient (CSQ). Like the OSQ, a value of one indicates symmetry within the phase and as the value increases, the opening segment is proportionally longer than the closing segment, which may indicate less tension in the vocal folds. Chen et. al report CSQ to be greater in women than in men. For detailed descriptions of how to determine time measures in an EGG wave and for reported numerical values for males and females please refer to Table 3.

In summary, males have greater airflow amplitude values than females. In terms of proportions of the period of a vibratory cycle, female vocal folds are open longer than male vocal folds. Further, females have the more symmetrical opening-closing segment durations during the open phase but males are more symmetrical during closed phase. Data from separate studies suggest that in a female glottal cycle, the majority of the closing time (b+c in Figure 1) is during the open phase (bSegment), whereas for males it is during the closed phase (cSegment). These data were collected separately using different protocols, and therefore they may not be directly comparable, however they do suggest that the greatest difference in quotient length between males and females are in the bQuotient and cQuotient. In other words, because more of the vocal fold closing movement occurs during the open phase instead of the closed phase, females have greater open quotient than males and air is flowing for a greater portion of the cycle.

Glottal Source Characteristics in Transgender Speakers

Research is needed to investigate glottal source characteristics of TG speakers to determine if and if so, how, TG glottal source characteristics are significantly different from the cisgender normative data. These data will either support the typical TG hypothesis and achieve or approximate the desired gender values or they will differ from both gender groups in an entirely new pattern. If the normative values are found to be different in either pattern, than the perceptual effects of these differences should be evaluated to determine if, and to what extent, glottal source characteristics should be targeted in the feminization of TG voice. Data from previous glottal source characteristic studies suggest that the MTF speaker values for airflow will not support the typical TG hypothesis.

Gorham-Rowan and Morris (2006) and Holmberg et al. (2009) studied some of these glottal source characteristics in TG voice. Holmberg et al. examines the relationship between self-rating, phonetograms, perceptual ratings, airflow, and pressure from 25 MTF speakers. This study found that while MTF derived subglottal pressure values remain within the normal limits for both male and female speakers, the data were greater for TG speakers than either gender and most similar to that of a cisgender male speaking in a high voice. This data may indicate that TG speakers are not using subglottal pressure to increase loudness, as is the traditional assumption, but to create a higher pitched voice. Average mean flow for the MTF speakers also was higher than either female or male normative data.

Gorham-Rowan and Morris examined the relationship between measures of airflow, F_0, and listener perceptions of vowel samples collected from 13 MTF speakers. Results show that MTF speakers had a higher maximum flow declination rate (MFDR) value and a higher DC flow value than either cisgender male or female speakers, indicating that MTF speakers close their vocal folds more quickly than, but not as completely as, cisgender male or female speakers. This study, however, only included airflow measures. Little can be inferred about timing measures or the closed portion of the vocal fold vibration cycle. Higher declination rates may indicate that TG speakers close their vocal folds more quickly than either male or female speakers, however, it is unclear if that would occur by decreasing the duration of the bSegment, cSegment, or both segments. Supplementing airflow amplitude data with simultaneous temporal measures of vocal fold oscillation may more illuminate the results found in the few glottal source characteristic studies conducted.

The purpose of this study is to gather pressure measures, air flow measures (peak flow, AC flow, DC flow, OQ, CQ, OSQ, aQuotient, and bQuotient), and EGG measures (cQuotient, dQuotient, CSQ) for MTF, cisgender male, and cisgender female speakers during structured speech tasks and to answer the following research questions:

1. Are there differences between MTF and cisgender males and females for derived peak subglottal pressure?
2. Are there differences between MTF and cisgender males and females in the amplitude measures of the airflow characteristics?
    a. Peak flow
    b. AC flow
    c. DC flow
3. Are there differences between MTF and cisgender males and females in cycle duration measures used to quantify the glottal source patterns?

a. open and closed quotients (i.e., phase duration in proportion to the period)
b. speed quotients (i.e., opening/closing duration ratio within a phase)
c. quarter quotients (i.e., four opening and closing durations in proportion to the period)

If there are significant differences between male and female speakers, but not between any other pair, that will indicate that the values for TG speakers are approximating or achieving the desired gender's values. If there are significant differences between the TG speakers and either of the cisgender groups, it will indicate that the TG speakers are using a vocal fold vibration pattern that does not conform to the typical TG hypothesis.

MTF airflow values could support the typical TG hypothesis and approximate or achieve female values or they could do something entirely different from either gender as compensation for using a biologically male larynx to create a female voice. Two airflow amplitude measures (e.g., AC flow and peak flow) are reported to be greater in males than in females, and the other, DC flow, is reported to be greater in females than in males. Following the trend of the typical TG voice hypotheses for acoustic measures, one may hypothesize that these airflow amplitude measures of MTF voice will approximate or achieve female values (i.e. decrease peak flow and AC flow and increase DC flow). This hypothesis seems less likely when considered from the perspective of previously published literature and unchanged anatomy. Gorham-Rowan and Morris (2006) find that these amplitude values do not conform to the typical TG hypothesis and shift away from female values. Additionally, a biologically male larynx may be physiologically unable to approximate or achieve the amplitude values of the desired gender. Therefore, it is hypothesized that airflow measures may not conform to the typical TG hypothesis and may not approximate or achieve values of the desired gender. Timing measures gathered from EGG and airflow may be able to explain why or how either of these options occur.

Airflow glottogram and EGG are most appropriate for measuring timing of phonation physiology and can describe a glottal cycle in its entirety when gathered at the same point in time. The literature shows that males have a larger closed quotient and a smaller open quotient than females. Within the quotients, published values for speed quotients indicate that females have a more symmetrical open speed quotient than males (a~b in Figure 1) and that males have a more symmetrical closed speed quotient than females (c~d in Figure 1). This will result in a smaller bQuotient and larger cQuotient for males compared to females. Based on this literature and the typical TG hypothesis, it is hypothesized that MTF speaker values will shift towards female speaker values by increasing bQuotient and decreasing cQuotient. The possibility remains, however, that TG speakers create an entirely new and different pattern that does not approximate desired gender values to create a feminine voice from a male larynx.

Overall, by examining the glottal source characteristics of MTF voice in comparison to male and female voice this study aims to explain how an MTF speaker modifies a variety of aspects of phonation to attempt feminine voice qualities. It is predicted that MTF glottal phase patterns will approximate cisgender female values but be distinct from either male or female cisgender values for airflow patterns. It is likely that if one set of values (amplitude/timing) approximates the desired gender that the other set of values will create a new pattern. These hypotheses are based on previously published data as well as knowledge about the physiology of the vocal folds.

Chapter 2: Methods

Participants

A convenience sample of 30 speakers participated in this study. Ten MTF speakers were recruited from a pool of previous research participants and clients of The George Washington University Speech and Hearing Center (GWUSHC) and word-of-mouth. Ten male and 10 female cisgender speakers were recruited through fliers and advertisements sent through the local community as well as word-of-mouth. All participants followed informed consent procedures as approved by the Institutional Review Board. FTM speakers were also recruited, however, only data for three participants was collected during this time.

To be included in the study, TG speakers had to present as their desired gender 100% of the time, and must have done so for at least one year. Cisgender speakers were required to have no history of gender deviant identity disorders. Exclusion criteria common to all groups included neurological or voice disorders, non-native speakers of English, and current enrollment in voice therapy. All participants were asked their age, height, weight, current medications, and history of voice disorders or therapy. Every participant was age-matched within five years to a participant from the other two groups. Demographic data of speakers are presented in Table 4.

Instrumentation

This study collected three different types of data: derived subglottal pressure, airflow, and glottal contact. All subglottal pressure and airflow equipment required calibration prior to data collection for two reasons. The first is that the transducers and recording equipment are sensitive and record amplitude data (on the y-axis) in slightly different manners with each use dependent on factors such as drift. The second is that the data is recorded in arbitrary units (CSL units) and there needs to be a record of the relationship between these arbitrary units and units of flow and pressure (L/s or cmH_2O) for use in later signal interpretation. A Glottal Enterprises PC-1H calibrator was used to calibrate pressure and a Glottal Enterprises MCU-4 Pneumotach Calibration Unit was used to calibrate airflow. The electroglottograph did not require calibration because the measures employed were all temporal.

Pressure data collection was achieved using a Glottal Enterprises PT-25 pressure transducer and a 1.25 inch piece of 3 mm plastic tubing used to gather pressure data from within the mouth while allowing the lips to maintain a tight seal. Airflow data collection was achieved using a pneumotach to collect airflow data consisting of an airflow transducer, the Glottal Enterprises PT-2E, and a Glottal Enterprises MA-1L CV pneumotach mask. Both pressure data and airflow data were relayed using a Glottal Enterprises MS-110 Transducer and Analog Computer Data Interface. Glottal contact was collected using EL-2 electrodes attached to a hook and loop fastener strap that held the electrodes in place on the thyroid lamina. The electronic signals from the electrodes were relayed to a Glottal Enterprises EG2-PCx Professional Voice Input Interface.

All instrumentation output was connected to the KayPentax model 4500 interface and recorded in the Computerized Speech Lab (CSL) program. The CSL program allows acquisition of up to four channels from up to four sources at once, so it is ideal for collecting simultaneous data from multiple pieces of instrumentation. Waveview Pro 2.4 software was used to inverse-filter airflow signals. SoundSwell software was used to analyze airflow and EGG signals because it is designed for precise time and amplitude analysis. Statistical Program for Social Sciences (SPSS) was used for statistical analysis of the data. Brands and types of instrumentation used were chosen for consistency with Gorham-Rowan and Morris (2006), Chen et al. (2002), and Higgins and Saxman (1991).

Calibration and Data Collection Procedure

Calibration of pressure and flow transducers.

Pressure and airflow transducers require calibration and zeroing to permit interpretation of the electronic signals in cmH_2O for pressure and L/s for airflow and to eliminate drift of the transducer signals. All transducers were calibrated within thirty minutes of data collection. Sampling rate in CSL was set to 22050 Hz. Calibration was achieved by zeroing both transducer signals in CSL by adjusting gain and offset. The zeroed transducers were then calibrated with known values by the calibration units. The pressure calibration unit applied pressure to the transducer in the amount of 0, 5, and 10 cmH_2O. The pneumotach calibration unit produced airflow rates in the amount of positive and negative 0, .5, and 1 L/s. These values were recorded and used to derive a calibration factor to be applied to participant data during analysis. Please see Appendix A for exact calibration procedures.

Data collection.

Data calibration and collection took place in a soundproof booth at the GWUSHC to prevent external noise from being recorded along with audio signals and given consent paperwork which was explained orally and in writing. After participants were given the opportunity to ask questions, the consent forms were signed. The examiner asked questions regarding demographic data and the speaker’s participation eligibility, specifically determining age, height, weight, current medications, desired gender, birth gender, and history of neurological, psychological, or physiological disorders. If all criteria were met, data collection was initiated. All participants were seated and a headset microphone was placed behind their ears and over their neck.

Data collection involved two phases with somewhat different instrumentation used for each phase. The initial phase involved collection of oral pressure estimates of subglottal pressure, and airflow signals during production a string of /pae/ syllables. The second phase involved collection of the oral airflow signal while the subject produced a sustained vowel. Each portion will be described separately below.

Repeated syllable phase.

To gather data for derived subglottal pressure, participants were asked to repeat the syllable /pæ/ as it was recorded with the pneumotach, pressure transducers, and a microphone. The syllable */pæ/* was chosen because it is contains the phoneme /p/ that is most likely to create an airtight seal despite pressing a tube between the lips (Smitheran & Hixon, 1981). Plastic tubing was attached to the oral pressure transducer and placed through the pneumotach mask so that the pressure that was recorded was the pressure behind the lips while closed. Participants were instructed to hold only the mask handle while pressing the mask against their face to avoid air leakage, and to position the end tip of the tubing between their teeth and in the center of their mouth without biting the tubing. Participants were instructed to repeat the syllable /pæ/ at a rate of twenty repetitions in seven seconds. The anchor word “pack” and the appropriate rate were demonstrated by the examiner along with the instructions to use the participant’s typical speaking voice and to keep rate and volume consistent. If the produced sample proved consistent and reliable (i.e. aspiration was visible in the airflow signal and clear peaks were present for pressure), the audio, airflow, and pressure signals were recorded.

Sustained vowel phase.

To gather data for the amplitude and duration measures, participants were asked to sustain the vowel /□/. The vowel /□/ was chosen so that data could be compared with relative ease across several studies. After acceptable productions of /pæ/ were elicited, participants were instructed to remove the mask from their face. The pressure transducer was removed from the mask and replaced with a small stopper. The EGG electrodes were placed on the participant's thyroid lamina and held in place while the participant produced test iterations of /□/ to determine if the electrodes were appropriately placed. Appropriate placement was determined by the Professional Voice Input Interface which indicates with a series of lights whether placement is too high or too low to accurately determine conductance. When conductance was determined to be within an acceptable range, the electrodes were strapped into place using a hook and eye strap. Speakers were asked to sustain a vowel for approximately 5 seconds three times with a several second pause for breath in between each production. Speakers were asked to use the modal register (i.e. speaking voice) and to keep volume consistent across all repetitions. If this was not achieved with these instructions, participants were coached with an anchor word (i.e. father) that contained the target vowel in order to orient the participant to how it is produced in typical speech. Speakers were given hand cues to start, stop, and breathe during their production. Acceptable productions (defined as visibly consistent airflow and EGG recordings) were recorded for audio, airflow, and EGG. See Appendix B for full descriptions of the speaker protocol.

Data Measurement

Repeated syllable analysis.

The repeated syllables were analyzed in CSL for peak derived subglottal pressure. The calibration factor was determined by averaging five data points from the 10 cmH_2O calibration data and subtracting drift. Calibration data were checked for accuracy by ensuring that the ratio of the 5 cmH_2O data and the 10 cmH_2O data approximated a value of one half. Five consistent peaks were selected from the recorded /pæ/ data. The five continuous peaks were chosen based on two criteria. First, the chosen peaks needed to be consistent with each other without any outliers included. Second, the chosen peaks needed to be representative of the majority of the iterations. Five consistent peaks were not selected if they appeared to be markedly different from other peaks in the selection. Once chosen and averaged, the resulting values were converted from CSL units to cmH_20 by multiplying the calibration factor. Please see Appendix C for full descriptions of calculating pressure values.

Sustained vowel analysis.

The calibration factor was determined by averaging fifty milliseconds of the 1 L/s calibration data and subtracting drift. Calibration data were checked for accuracy by ensuring that the ratio of the .5 L/s data and the 1 L/s data approximated a value of one half. Individual airflow data were analyzed using CSL, Waveview Pro 2.4 and SoundSwell software. Individual airflow data were analyzed by isolating the airflow and EGG channels from the audio channel for prolonged vowel /*a*/. The isolated channels were separated into the three elicited tokens and the second token was entered in Waveview Pro 2.4. This token was inverse filtered beginning 2.08 seconds from the beginning of the sample, continuing for 10-15 cycles (Rothenberg, 1972). After the experimenter was trained in the process of inverse filtering, all signals were filtered by a single experimenter to ensure that all signals were filtered in a similar manner for reliability. Inverse filtering involves filtering out resonances of the vocal tract from the waveform. This involves initially adjusting three notch filters for the vowel based on the gender of the speaker (Peterson and Barney, 1952) and setting the time constant to .30 ms. Additionally a low pass filter was set to 2300 Hz or above to eliminate the influence of higher frequency resonances. Filters were then more carefully adjusted from these published frequencies, beginning with the lowest frequency filter for general shape, then the second filter to smooth out higher frequency resonances, and ending with the third filter to smooth the wave. Filtered files were subsequently measured in SoundSwell.

Time and amplitude were determined for the peak and valley for each cycle by zooming in as close as possible to the selected waves and then making a series of measures as depicted in Figure 1. Initially the cursor was adjusted to determine the highest and lowest points on the selected wave (amplitude and time at T_1 and T_3 on Figure 1). Fifteen percent of the difference between average of peak values and average of valley values was calculated to determine the boundary line between the open and closed phases of the cycle (Rothenberg & Mahshie, 1988). Time values were recorded for any point in the selected waves crossing this line (T_0 and T_2 on Figure 1). Amplitude-based values from the peak and the valley were used to calculate peak flow (amplitude at T_1 on the flow glottogram multiplied by the calibration factor), DC flow (amplitude at T_3 on the flow glottogram multiplied by the calibration factor), and AC flow (amplitude at T_1 minus amplitude at T_3 multiplied by the calibration factor). Time-based values were used to calculate open quotient (T_2 minus T_0 divided by the cycle's duration), closed quotient (T_0 of the next cycle minus T_2 of the cycle divided by the cycle's duration), open speed quotient (T_1-T_0 divided by T_2-T_1 on the flow glottogram), aQuotient (T_1-T_0 divided by the cycle's duration) and bQuotient (T_2-T_1 divided by the cycle's duration) (*see* Tables 1 and 2 for definitions and descriptions of calculations and Figure 1 for visual representation of T_0, $T_{1,}$ T_2, and T_3).

Airflow and EGG signals were temporally aligned so that measures from airflow and EGG signals reflected the same glottal cycle. As with the airflow signals, a boundary line was calculated between the open and closed portions of the cycle at 15% of the difference between the average peak and valley amplitudes. The selected waves were opened in SoundSwell and the zoom was adjusted as close as possible to the selected wave. Time was recorded for the peak (T_3 on Figure 1), valley (T_1 on Figure 1), and each point where the boundary line crossed the selected EGG cycles (T_0 and T_2 on Figure 1). Time-based values were used to calculate closed speed quotient (T_0 of the next cycle-T_3 divided by T_3-T_2), cQuotient (T_3-T_2 divided by the cycle's duration) and dQuotient (T_0 of the next cycle-T_3 divided by the cycle's duration) (*see* Table 3 for definitions and Figure 1 for visual representations of T_0, $T_{1,}$ T_2, and T_3).

Statistical analysis.

Group data were entered into SPSS using Bonferroni correction. Descriptive statistics were generated and an analysis of variance (ANOVA) was calculated between groups for subglottal pressure. A multivariate analysis of variance (MANOVA) was calculated between groups for peak pressure, DC flow, and AC flow. A multivariate analysis of variance (MANOVA) was calculated between groups for open quotient and closed quotient, for open speed quotient and closed speed quotient, and for all quarter quotients.

Chapter 3: Results

Subglottal pressure values were presumed equal to the greatest intraoral pressure values immediately prior to the plosive release in /pæ/. Males had a mean subglottal pressure value of 4.967 cmH_2O (1.22), females had a mean value of 5.332 cmH_2O (1.923) and MTF speakers had a mean value of 6.195 cmH_2O (1.36) represented in Figure 2 and Table 5. Although differences between these gender groups did not reach statistically significant levels in an ANOVA ($F(2,25)=1.634$, $p=.215$), represented in Table 7, MTF speakers had the greatest derived peak subglottal pressure, followed by females and then males with the smallest value.

Peak flow is the greatest rate of airflow during a cycle and DC flow is the least rate of airflow during a cycle and usually represents leakage during the closed phase. Together, they can be used to calculate AC flow, which represents the airflow that is influenced by the vocal folds opening and closing. Females had a mean peak flow value of .215 L/s (.077), TG speakers had a mean peak flow value of .347 L/s (.141) and males had a mean peak flow value of .449 L/s (.136). Females had a mean DC flow value of .052 L/s (.072), TG speakers had a mean DC flow value of .083 L/s (.080) and males had a mean DC flow value of .111 L/s (.056). Females had a mean AC flow value of .162 L/s (.049), TG speakers had a mean AC flow value of .264 L/s (.087) and males had a mean AC flow value of .339 L/s (.124). All values are represented in graph and table form in Figure 3 and Table 5. A MANOVA revealed a significant difference among groups for AC flow and peak flow, but not for DC flow (see Table 7). Pairwise comparisons reveal significant differences between male and female groups for Peak flow ($p=.001$) and AC flow ($p=.001$), with MTF values between these two groups not significantly different from either group (see Table 9). Male voice has the greatest rate of airflow for peak flow, DC flow, and AC flow compared to MTF and female speakers, with MTF speaker values between the cisgender values. This supports the hypothesis that MTF speakers are achieving values more similar to those of the desired gender, yet still between male and female cisgender values.

A second MANOVA tested differences between groups for all the duration-based measures. The open and closed quotients are the percentage of time that the vocal folds are open or closed for the total duration of the cycle. Male speakers had an average open quotient of .547 (.096), female speakers had an average open quotient of .65 (.086), and TG speakers had an average open quotient of .721 (.061). Transgender speakers had an average closed quotient of.279 (.061), female speakers had an average closed quotient of .35 (.086), and male speakers had an average closed quotient of .453 (.096). These values are represented graphically in Figure 4 and in table form in Table 7. There was a significant difference between groups for both open and closed quotient as is presented in Table 8. Pairwise comparisons reveal significant differences between male and MTF groups for open quotient and closed quotient ($p \leq .000$ for both quotients) and between male and female groups for open quotient and closed quotient ($p=.028$ for both quotients), but not between female and MTF groups ($p=.193$ for both quotients). The MTF group had the greatest open quotient value and the smallest closed quotient value. This indicates that open and closed quotients of MTF speakers do not support the hypothesis that their values would be between those of the cisgender groups; rather, MTF and cisgender female values are statistically equivalent, with the female values slightly closer to the males.

Open and closed speed quotients describe the relative symmetry within the open and closed portions of the cycle. A quotient close to a value of one will represent a phase that has symmetrical segments, or the duration of the opening and closing segments will be approximately the same amount of time. A quotient larger than one will have a less symmetrical phase with a greater amount of time allotted to the opening segment of the phase rather than the closing segment of the phase. Transgender speaker CSQ values had an average of 2.149 (1.022), male speaker CSQ values had an average of 2.273 (.954) and female speaker CSQ values had an average of 2.395 (1.97). Male speaker OSQ values had an average of 1.395 (.27), TG speaker OSQ values had an average of 1.548 (.556), and female speaker OSQ values had an average of 1.601 (.372). All values are represented in graph and table form in Figure 5 and Table 6. Although between group differences did not reach statistically significant levels for OSQ ($F(2,27)=.659$, $p=.526$) or CSQ ($F(2,27)=.078$, $p=.925$), represented in Table 8, female speakers had the greatest open and closed speed quotients (i.e., least symmetrical and toward a longer opening segment) followed by MTF and then males with the smallest values in OSQ and followed by males then MTF in CSQ. Speed quotient value differences were small and statistically insignificant.

The quarter quotients (i.e., aQuotient, bQuotient, cQuotient, and dQuotient) describe the proportion of each quarter to the duration of the total cycle. A perfectly sinusoidal wave would have four quarter quotient values of .25. Male speakers had an average aQuotient, or the opening of the open phase, value of .313 (.055), female speakers had an average aQuotient value of .392 (.099), and TG speakers had an average aQuotient value of .423 (.079). Male speakers had an average bQuotient, or the closing of the open phase, value of .234 (.055), female speakers had an average bQuotient value of .267 (.059), and TG speakers had an average bQuotient value of .297 (.065). Female speakers had an average cQuotient, or closing of the closed phase, value of .196 (.080), male speakers had an average cQuotient value of .216 (.065), and TG speakers had an average cQuotient value of .221 (.090). Transgender speakers had an average dQuotient, or the opening of the closed phase, value of .399 (.081), female speakers had an average dQuotient value of .445 (.078), and TG speakers had an average dQuotient value of .454 (.095). The values are represented in graph and table form in Figure 6 and Table 6. There was a significant difference between groups for aQuotient ($F(2,27)=5.117$, $p=.013$), but not for bQuotient ($F(2,27)=2.764$, $p=.081$), cQuotient($F(2,27)=.293$, $p=.749$), or dQuotient ($F(2,27)=1.192$, $p=.319$), represented in Table 7. Pairwise comparisons for aQuotient revealed significant differences between the MTF and male groups ($p=.013$), but female values were not significantly different from MTF ($p=1.0$) or male ($p=.103$) speaker groups, represented in Table 9. Like with the open and closed quotients, the MTF speaker values not only shifted significantly away from male values toward female values, they slightly surpass female values for aQuotient. Although not significant, this pattern was the same for bQuotient and dQuotient, but not cQuotient. It is unclear how b, c, and d Quotients are modified to compensate when the biologically male larynx increases aQuotient, but examination of the non-significant mean scores suggests it is possible that MTF speakers decrease dQuotient in order to increase aQuotient.

In summary, the large-scale MTF vocal fold opening and closing patterns (i.e., open and closed quotients) have changed from the biologically male patterns and are now most similar to those of a female, even slightly surpassing them. The smaller quarter quotient adjustments made to achieve this pattern are most apparent in the aQuotient, when the vocal folds are first beginning to open. Additionally, subglottal pressure and airflow measures of MTF voice seem to shift toward female values (pressure may have even surpassed female values) but not to the extent that they are statistically different from male values.

Chapter 4: Discussion

Interpretation of the Data

The typical hypothesis in transgender research is that as values move toward the normative values of the desired gender, the likelihood that the TG individual is accepted as the desired gender increases. The results of this study suggest that this hypothesis may not be the case for all measures and that suggestion has consequences for all aspects of TG voice feminization not currently supported by specific TG research. The current study's findings suggest that MTF speakers alter their vocal fold vibration pattern (duration measures) into something different from either gender to achieve airflow amplitude values that approximate those of the desired gender.

This new pattern was made up of an increased open quotient created by an increased aQuotient. It is of limited use to look at measures that did not distinguish male from female groups (i.e., subglottal pressure, DC flow, both SQ and the three nonsignificant quarter quotients); however, examination of the mean values reveals the MTF group, while not significantly different in these measures, was between the male and female values for two measures (DC flow, OSQ), farthest from the male group in four measures (subglottal pressure, a, b, and dQuotients), and farthest from the female group in two measures (CSQ, cQuotient). Only by examining a variety of measures of glottal behavior can it be seen that it is most likely that MTF speakers did not conform to the typical TG hypothesis; rather, MTF speakers create a different pattern within the timing aspects of phonation in order to approximate desired gender values in amplitude measures.

Interpretation of Data within the Context of Previously Published Results

This study's data are in many ways consistent with previously published results. Perkell et al. (1994) report peak flow for male and female speakers to be .410 L/s and .286 L/s while this study found peak flow for male and female speakers to be .449 L/s and .215 L/s. They report AC flow for males and females as .331 L/s and .156 L/s while this study found AC flow to be .339 L/s and .162 L/s respectively. The values for DC flow were not significantly different between genders in either study. Higgins and Saxman (1991) find the male open and closed quotient to be 57% and 50.8% respectively and the female open and closed quotient to be 63% and 44.5% respectively. This study finds that the male open and closed quotient to be 54.7% and 45.3% respectively and the female open and closed quotient to be 65% and 35% respectively. Higgins and Saxman also find that the bQuotient and OSQ for males is 19.7% and 1.52 and for females is 23.7% and 1.36. This study finds the bQuotient and OSQ for males to be 23.4% and 1.39 and for females to be 26.6% and 1.6. While OSQ values reversed for males and females, the gender differences were not statistically significant in either study. Finally, Chen et. al (2002) find cQuotient and dQuotient to be greater in males than in females. Exact values are not comparable because Chen et al. reported in milliseconds and not in quotient format. They also find that CSQ is 1.39 for males and 2.09 for females, much like this study's 2.27 for males and 2.39 for females. Overall, this study's values were consistent with previously published cisgender results.

The airflow values of the MTF speakers in this study were unlike those of the MTF speakers in the Gorham-Rowan and Morris (2005) study, which reported MTF to have greater airflow than either gender group. It was originally hypothesized that MTF speakers make adjustments within the closing segment of the vocal fold vibration to increase their amplitude measures beyond those of either gender. This study's results were not consistent with Gorham-Rowan and Morris because the data suggest that TG speakers do not increase amplitude values beyond those of either gender, but approximate their values to their desired gender's values. The results of this study suggest the opposite of the hypothesis that MTF speakers would adjust the closing phase (b+c) and instead suggests that TG speakers modify the open phase (a+b) of the glottal cycle to create airflow amplitude measures that approximate those of the desired gender.

The current study's data further suggest that the MTF speaker increases the amount of time that the glottis is open by increasing the duration of time that it is opening during the open phase (i.e., aQuotient). In other words, the MTF speaker increases the aQuotient in order to achieve a more feminine voice instead of increasing the bQuotient to approximate female values. No other quarter quotient value for MTF speakers was significantly different from males, so it remains unclear which quotient(s) decreased to allow for an increased aQuotient. Nonsignificant data trended toward a larger bQuotient for MTF speakers, like in cisgender females. It does not appear, however, that cQuotient was decreased, as was originally hypothesized, to give greater length to aQuotient and bQuotient. Instead, it appears that dQuotient may have decreased to add length to the aQuotient. Additional data should be collected to determine if these trend-level findings may be significant with a larger sample and more statistical power.

While differences in derived peak subglottal pressure were not significant between groups, the data are greater for MTF speakers than for either cisgender group. MTF speakers had derived peak subglottal pressure values slightly greater than either male or female speakers, who had almost identical values. This increase in subglottal pressure could explain the smaller dQuotient (i.e., faster opening to end the closed phase) because a greater amount of subglottal pressure would increase the speed at which the vocal folds open and close. The dQuotient represents the portion of the cycle in which the vocal folds are transitioning from fully closed to beginning to open. The source-filter theory states that subglottal pressure is what forces the vocal folds to open, therefore the subglottal pressure influences the dQuotient in particular. This subglottal pressure may not be used, as it is thought traditionally, to increase tension and therefore loudness, but to change the vocal fold vibration pattern while using loose vocal folds. Speed quotient data, while not significant, support this potential explanation because MTF had an open speed quotient more similar to females than to males with MTF having a smaller closed speed quotient than either gender. Although not significant, the closed phase was most symmetrical for MTF speakers, supporting the interpretation that dQuotient is the segment that decreases in MTF speakers to compensate for a longer aQuotient.

Limitations

The first limitation of this study was the availability of participants. The MTF participants were fairly easy to recruit as they were typically past clients of the GWUSHC. The inclusion of FTM participants in statistical calculations could better inform whether TG speakers adjust glottal characteristics toward those of their desired gender's normative values, and if FTM and MTF adjust in similar patterns (i.e., partially to be between cisgender values, or equivalent to desired gender, or perhaps beyond desired gender in a unique pattern). Increasing the number of participants would increase the statistical power of the results. The second limitation was the format in which data were collected. The calibration factor was applied after data were collected and this limited the type of measures that could be calculated. The third limitation is that the MTF speaker group in this study may have included speakers who would not be perceived as female to listeners. However, all MTF participants had received voice feminization therapy..

Future Directions

Future studies should continue to investigate glottal physiology of the TG larynx with both visualization and perceptual data. It is important to consider the three dimensional nature of the larynx. The current study’s examination of mucosal wave to quantify the timing and phase shifts of the glottal cycle could be supplemented by visual examination of glottal configuration (e.g., posterior chink). No conclusions can be drawn about glottis size without supplementing aerodynamic measures with visualization data. This study’s data are limited to describing how MTF speakers create a female voice by altering timing of the mucosal wave, however future studies could include visualization data to elaborate upon these conclusions.

All, some, or none of the statistically significant values (peak flow, AC flow, aQuotient, open quotient, or closed quotient) may relate to the perception of gender. If any of these values are found to correlate with the perception of gender in voice, then it might be important for guiding assessment and treatment of TG voice. The current study’s data are informative and suggest that the MTF laryngeal physiology does not approximate female values, which could lead to a clinical practice paradigm shift away from the current approach of simply aiming for the desired gender’s normative value.

References

Adler, R., Hirsch, S., & Mourdaunt, M. (Eds.). (2006). *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide*. San Diego, CA: Plural Publishing.

American Psychological Association. (2000). *Diagnostic and statistical manual of mental disorders: Text revision* (4th ed.). Arlington, VA: American Psychiatric Association.

American Speech-Language-Hearing Association. (2007). *Scope of practice in speech-language pathology* [Scope of Practice]. Retrieved March 4, 2011, available from www.asha.org/policy.

American Speech-Language-Hearing Association. (2011). *Voice and communication therapy for TG/TS voice*. Retrieved March 4, 2011, available from www.asha.org/public/speech/disorders/TGTS.htm.

Andrews, M. L. (2006). *Manual of voice treatment: Pediatrics through geriatrics* (3rd ed. ed.). Clifton Park, NY: Thomson Delmar Learning.

Boone, D., McFarlane, S., Von Berg, S., & Zraick, R. (2010*). The voice and voice therapy (8th ed.)*. Boston, MA: Allyn & Bacon.

Carew, L., Dacakis, G., & Oates, J. (2007). The effectiveness of oral resonance therapy on the perception of femininity in male-to-female transsexuals. *Journal of Voice, 21(*5), 591-603.

Chen, Y., Robb, M., & Gilbert, H. (2002). Electroglottographic evaluation of gender and vowel effects during modal and vocal fry phonation. *Journal of Speech, Language, and Hearing Research, 45*, 821-829.

Childers, D., & Wu, K. (1991). Gender recognition from speech. *Journal of Acoustical Society of America, 90*(4), 1841-1856.

Coleman, R. (1983). Acoustic correlates of speaker sex identification: Implications for the transsexual voice. *The Journal of Sex Research*, *19*(3), 293-295.

Dacakis, G. (2000). Long-term maintenance of fundamental frequency increases in male-to-female transsexuals. *Journal of Voice*, *21*(4), 549-556.

Ferrand, C. (2007). *Speech science: An integrated approach to theory and clinical practice.* Boston, MA: Allyn & Bacon.

Gelfer, M. (1999). Voice treatment for the male-to-female transgendered client. *American Journal of Speech-Language Pathology, 8*, 201-208.

Gelfer, M., & Mikos, V. (2005). The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. *Journal of Voice, 19*(4), 544-554.

Gelfer, M., & Schofield, K. (2000). Comparison of acoustic and perceptual measures of voice in male-to-female transsexuals perceived as female versus those perceived as male. *Journal of Voice*, *14*(1), 22-33.

Gorham-Rowan, M., & Morris, R. (2006). Aerodynamic of male-to-female transgender voice. *Journal of Voice*, *20*(2), 251-262.

Gunzburger, D. (1995). Acoustic and perceptual implications of transsexual voice. *Archives of Sexual Behavior, 24*(3), 339-347.

Higgins, M., & Saxman, J. (1991). A comparison of selected phonatory behaviors of healthy aged and young adults. *Journal of Speech and Hearing Research, 34*, 1000-1010.

Hillenbrand, J., & Clark, M. (2009). The role of f_0 and formant frequencies in distinguishing the voices of men and women. *Attention, Perception & Psychophysics, 71*(5), 1150-1166.

Hixon, T., Weismer, G., & Hoit, J. (2008). *Preclinical speech science: Anatomy physiology acoustics perception*. San Diego, CA: Plural Publishing.

Holmberg, E., Hillman, R., Perkell, J., & Gress, C. (1994). Relationships between intra-speaker variation in aerodynamic measures of voice production and variation in SPL across repeated recordings. *Journal of Speech and Hearing Research, 37*, 484-495.

Holmberg, E., Oates, J., Dacakis, G., & Grant, C. (2009). Phonetograms, aerodynamic measurements, self-evaluations, and auditory perceptual ratings of male-to-female transsexual voice. *Journal of Voice, 24*(5), 511-522.

Kent, R. (1997). *The speech sciences*. Clifton Park, NY: Thomson Delmar Learning.

Klatt, D.H. & Klatt, L.C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*(2), 820-857.

Leaper, C., & Ayres, M. (2007). Review of gender variations in adults' language use: Talkativeness, affiliative speech, and assertive speech. *Personality and Social Psychology Review, 11*, 328-364.

Linville, S. (1992). Glottal gap configurations in two age groups of women. *Journal of Speech and Hearing Disorders, 35*, 1209-1215.

McCrea, C. (2004). The effects of vocal training, phonatory task, and gender on voice onset time. Unpublished Doctor of Philosophy, Florida State University, Tallahassee, FL.

Melcon, M., Hoit, J., & Hixon, T. (1989). Age and laryngeal airway resistance during vowel production. *Journal of Speech and Hearing Disorders*, *54*, 282-286.

Mount, K., & Salmon, S. (1988). Changing the vocal characteristics of a postoperative transsexual patient: A longitudinal study. *Journal of Communication Disorders, 21*(3), 229-238.

Oates, J., & Dacakis, G. (1983). Speech pathology considerations in the management of transsexualism: A review. *British Journal of Disorders of Communication: The Journal of the College of Speech Therapists in London, 18*(3), 139.

Rothenberg, M. (1973). A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *Journal of Acoustical Society of America, 53*(6), 1632-1645.

Rothenberg, M. (1977). Measurement of airflow in speech. *Journal of Speech and Hearing Research, 20*, 155-176.

Rothenberg, M., & Mahshie, J. (1988). Monitoring vocal fold abduction through vocal fold contact area. *Journal of Speech, Language, and Hearing Research, 31*, 338-351.

Sapienza, C., Stathopoulos, E. & Dromey, C. (1998). Approximations of open quotient and speed quotient for glottal airflow and EGG waveforms: Effects of measurement criteria and sound pressure level. *Journal of Voice, 12*(1), 31-43.

Smitheran, J., & Hixon, T. (1981). A clinical method for estimating laryngeal airway resistance during vowel production. *Journal of Speech and Hearing Disorders,* *81*, 138-147.

Sodersten, M., & Lindestad, P. (1990). Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech, Language, and Hearing Research, 33,* 601-611.

Spencer, L. E. (1988). *Speech characteristics of male-to-female transsexuals: A perceptual and acoustic study.* Folia Phoniatrica, 40, 31-42.

Sulter, A., & Wit, H. (1996). Glottal volume velocity waveform characteristics in subjects with and without vocal training, related to gender, sound intensity, fundamental frequency, and age. *Journal of Acoustical Society of America, 100*(5), 3360-3373.

Wolfe, V., & Ratusnik, D. (1988). Acoustic and perceptual measurements of roughness influencing judgments of pitch. *Journal of Speech and Hearing Disorders, 53,* 15-22.

Wolfe, V., Ratusnik, D., Smith, F., & Northrop, G. (1990). Intonation and fundamental frequency in male-to-female transsexuals. *Journal of Speech and Hearing Disorders*, *55*, 43-50.

Appendix A

Calibration Procedures

Set-up.

1. Turn on the computer and log onto Hancock profile
2. Open the CSL main program
3. Turn on the CSL box
    a. "DC Coupled" buttons should be on (show green light) for Channels 3and4.
4. To set up CSL
    a. Go to options and then capture
    b. Check the boxes for channels 3 and 4 and make channel 3 active
    c. Set the sample rate at 22050 Hz
    d. Set the capture time for 15 seconds
    e. Verify that the input device is "CSL 4500 (ASIO)
    f. Close the scope window and draw the A window to be whole screen so you can see it.
5. To set up the MS110 box
    a. The power source is "EXT" and the power plug is plugged in.
    b. "Modulator" are both off
    c. "Input Invert" are both off
6. On the front of the MS110 box
    a. "TC" is off
    b. "Offset" black switches are on ADJ
    c. "Gain" black switches are on Fixed

d. "Input AandB" will be dual cord
e. The "A" cord should be attached to "PT-2E" Transducer (air flow)
f. The "B" cord should be attached to "P25" Transducer (pressure)

Calibrate for Airflow

1. Set up the pneumotach calibration unit
   a. Plug the power cord into the power source
   b. Turn the power on in the back
   c. Remove the black stopper from the top
2. Set up the mask
   a. Don't touch the mesh circles
   b. Wipe after each use
3. Calibrate for airflow
   a. Plug the PT-2E transducer into the mask
   b. Plug the other opening
   c. Ensure that mode on pneumotach calibration unit is on "flow"
   d. "Record" on CSL
      i. Adjust "offset" (Channel A) until green line disappears in center line
      ii. Since no flow is going into the mask, there should be no flow registering on MS110/CSL
   e. Place the mask on the gray base on top of the pneumotach calibration unit
   f. Turn the flow rate to 1/2 liter/sec and the volume to 1 liter
   g. Press the mask down onto the gray base

h. Press the start lever on the calibrator and record the signal until the calibrator stops

i. Save signal in Hancock>Documents>Study-TG Speech and Voice>Calibration> (Make folder labeled with participant ID)> half liter +Participant ID (ex: half liter SM001)

j. Turn the flow rate to 1 liter/sec and leave the volume at 1 liter

k. Press the mask down onto the gray base

l. Press the start level on the calibrator and record the signal until the calibrator stops

m. Save signal in Hancock>Documents>Study-TG Speech and Voice>Calibration> (Make folder labeled with participant ID)> liter+Participant ID (ex: liter SM001)

Calibrate for Pressure

1. Set up the PC-1H calibrator
    a. Pull up the syringe
2. Set up the mask
    a. Don't touch the mesh circles
    b. Wipe after each use
3. Calibrate for pressure
    a. Plug the P-25 transducer into the mask
    b. "Record" on CSL
        i. Adjust "offset" (Channel A) until green line disappears in center line

   ii. Since no pressure is going into the mask, there should be no pressure registering on MS110/CSL

c. Plug the P-25 transducer into the "transducer port" on the front of the PC-1H calibration unit

d. Push down the syringe until meter says 10 cmH_2O and record on CSL

e. After approximately 4 seconds, pull up the syringe so the meter says 5cmH_2O

f. After approximately 4 seconds, press the "pressure release" button until the gauge goes down to 0.

g. Save signal in Hancock>Documents>Study-TG Speech and Voice>Calibration> (Make folder labeled with participant ID)> pressure +Participant ID (Ex: pressure SM001)

Appendix B

Speaker Protocols

Step 1: Intro and consent.

"Thank you for coming to participate in this study! Right now we're studying some basic acoustic and aerodynamic physiology of voice, especially transgender voice. This will give our field a lot more information and help us develop healthy voice techniques most effective for changing how people perceive a transgender voice. I will be asking you some questions and then we'll get started with the voice part, ok?

First I need you to complete some paperwork. This form explains what the study is about and tells you about our procedures to keep everything confidential, potential risks and benefits to you, and that you will receive $10 when we finish here today. Please read it over and let me know if you have any questions."

Step 2: Questionnaire.

"I just have a couple questions to ask you…" Tester should write on questionnaire and ask questions verbally. Make sure you have the correct questionnaire (transgender vs non-transgender). If a participant does not meet (even one) criteria, cannot participate and therefore no payment.

Step 3: Derived subglottal pressure and flow rate.

Use the mask with pressure and flow transducers and a microphone. This procedure is modeled off of Holmberg. et al (2009). Instructor should say "Say /pæ pæ pæ / about 20 times or for approximately 7 seconds in a monotonous, smooth pattern. Keep your pitch, volume, and rate constant (3 syllables per second). Every /pa/ should sound about the same. Let's practice first." Instructor should change CSL Channel 4 to

EGG and place EGG transducers on person and test signal. Remove pressure transducer/tubing and replace with red plug in mask. In CSL, go to Options then Capture and change the time to 40 seconds

Step 4: Sustained vowels in modal and glottal register.

Use the mask with the flow transducer and a stopper in the second opening as well as the EGG transducers and a microphone for this task. This procedure is modeled off of Chen, Robb, and Gilbert (2002), Gelfer and Schofield (2000) and Gorham-Rowan and Morris (2006). Instructor says “Now I would like you to say the vowel three times and hold it out for about 6 seconds each time. Try to stay at the same volume so don’t go loud and soft. I will use my hand for stop and go, like this” (move hand back and forth for “go,” and hold up for “stop”). Repeat these instructions for /□/ /æ/ /u/ and /i/.

Appendix C

To Calculate Pressure Values

1. Open CSL from desktop then Main CSL Main Program button

2. Open calibration file in Window A

   a. Documents then Study-TG Voice and Speech then Calibration then filename will include participant ID number and "pressure" (after opening, will need to streth Window A to see all channels. Pressure is Channel 4)

   b. Scroll until red cursor in highest section in Channel 4 (this was when 10 cmH_2O were flowing) and check the y-axis value below in the bottom bar of the window.

   c. Average 3 representative points (ignore outliers) and enter this number into the Excel spreadsheet, "pressure" tab in the "calibration 10 cmH_2O" column.

   d. Put red cursor in lowest section, which may appear to coincide with the 0 zero (this was when 0 cmH2O were flowing) and check the y-axis value below in the gray area of the window.

   e. Average 3 representative points (ignore outliers) and enter this number into the Excel spreadsheet, "pressure" tab in the "calibration 0 cmH_2O" column.

3. Open papapa file in Window A (ok to remove calibration file away)

   a. Documents then Study-TG Voice and Speech then papapa

   b. Isolate Channel 4

i. Activate Channel 3 by clicking in it. Edit then multichannel function then split from next channel in source

ii. This will put Channel 4 in Window B. Activate Window B by clicking on it. You can stretch it larger.

iii. Save as SpeakerID papapa FP

c. Select the middle 5 peaks using the blue cursors. (insert comment on spreadsheet if one seems like an outlier)

d. View then Selected Data (Alt+V)

e. For *each* of those 5 peaks, put the red cursor on the peak (use the space bar to move slightly left and right) and record the y-value and time on the Excel spreadsheet in the "pressure" tab

f. Determine the average for time and y-value between each selected peak. Record the average for time and y-value on the Excel spreadsheet.

4. Before opening the next participant's files, you will need to reset

a. File then RESET (to the User Configuration)

Appendix D

Calculating Flow and EGG Data

To calculate flow values.

1. Inverse filter (WaveView Pro)
    a. Open file with "FEa" at end of filename (or FEb, FEc)
    b. Adjust Gain to x10, Zoom to 640ms, scroll to 2.08s
    c. Select vowel
    d. Turn on 50 Hz LPF
    e. Set the time adjustment for .30 ms for consistency.
    f. Do not set the high-pass filter lower than 2400 Hz.
    g. Turn on all filters, adjust according to first cycle beginning after 2.08s
        i. For general shape, adjust F_1 filter. Adjust F_1 until all large secondary peaks have disappeared. If large secondary peaks are present, F_1 is likely too high.
        ii. For smaller secondary speaks, adjust the F_2 filter.
        iii. For smoothing, adjust the F_3 filter.
        iv. If a file is aperiodic, skip cycles that will require filters more than 300 Hz different at anyone filter than filters set for previous cycles and restart filtering at the next periodic cycle.
    h. Save file, adding "filtered at {insert time using – instead of .} for {#} cycles. Ex: MTF001 mod1 FE filtered at 2-31 sec for 4 cycles
    i. Repeat process, beginning with the next cycle (so you will need to note time to start next file)

j. Do this for 10-15 cycles.

2. Calibration files (Swell)

a. Open file in CSL Model 4500

i. Documents then Study-TG Speech and Voice then Calibration

b. Expand Window A

c. Click on Channel 1

d. Edit then Multi-Channel functions then Split from next Channel in Source (now Flow data will be in the new “B” Window)

e. With Window B active, File then Save then (use same file name but save as .wav file)

f. File then RESET (to the User Configuration)

g. Repeat steps a through f

h. Transfer files to jump drive

i. Transfer files from jump drive to Flow Calibration data folder on Desktop

j. Open SoundSwell (Programs then Soundswell then Editor)

k. Open 1 liter file for participant (make sure you are looking for .wav files)

l. Select a section of 3 levels (positive, zero, negative) by left clicking and dragging, then Zoom (View menu or Ctrl+Z)

m. Highlight selection of positive signal to cover 30 seconds

n. Measure (View Menu or Shift+F5)

i. The t value should be near .5 (half minute) and the y-value is the average for that selection.

o. Record the y-value to the thousandths place in the spreadsheet

p. Repeat for zero section and negative section

q. Repeat for ½ liter file

3. Measure Flow (Swell)

   a. Open SoundSwell

   b. Open file (make sure you are looking for .wav files) from the filtered data folder on desktop

   c. Select the section near the time noted in the spreadsheet by left clicking and dragging, then Zoom (Edit menu or Ctrl+Z)

      i. May repeat several times until you see the cycles of interest

   d. Measure (View Menu or Shift+F5)

      i. The t value is the time (x axis) and the y-value is amplitude unit (y-axis)

   e. Use the right and left arrows to move the cursor to points T0, T1, T2, min, T3

   f. Calculating the 15% criterion line:

      i. 15% line = min + .15(max-min)

      ii. **T0 and T2 should cross at 15% line for both airflow and EGG signals

   g. The excel spreadsheet equations will fill the y-values for T0 and T2. Use the cursor to find this y-value in each phase (open and closed) of the cycle- when the cursor is at the y-value, that spot becomes T0 or T2.

      i. Enter the x-values for all T0 and T2.

      ii. Sometimes the cursor cannot land exactly on your desired y-value. Try to zoom in, and use the closet y-value to your desired y-value.

iii. Record T0 for next cycle if more than one inverse-filtered file is saved.

To calculate EGG values.

1. Measure EGG (Swell)
    a. Open SoundSwell Editor
    b. Opem the file from the EGG data folder on the desktop.
    c. Select the section near the time noted in the spreadsheet by left clicking and dragging.
    d. Zoom (Ctrl + Z)
        i. May repeat several times until you see the cycles of interest
    e. Measure (Shift + F5)
        i. The t value is the time (x-axis) and the y-value is amplitude unit (y-axis)
    f. Use the right and left arrows to move the cursor to points T0, T1, T2, min, T3.
    g. Calculating the 15% criterion line:
        i. 15% line = min + .15(max-min)
        ii. T0 and T2 should cross at 15% line for both airflow and EGG signals
    h. The excel spreadsheet equations will fill the y-values for T0 and T2. Use the cursor to find this y-value in each phase (open and closed) of the cycle- when the cursor is at the y-value, that spot becomes T0 or T2.
        i. Enter the x-values for all T0 and T2.

ii. Sometimes the cursor cannot land exactly on your desired y-value. Try to zoom in, and use the closet y-value to your desired y-value.

iii. Record T0 for next cycle if more than one inverse-filtered file is saved.

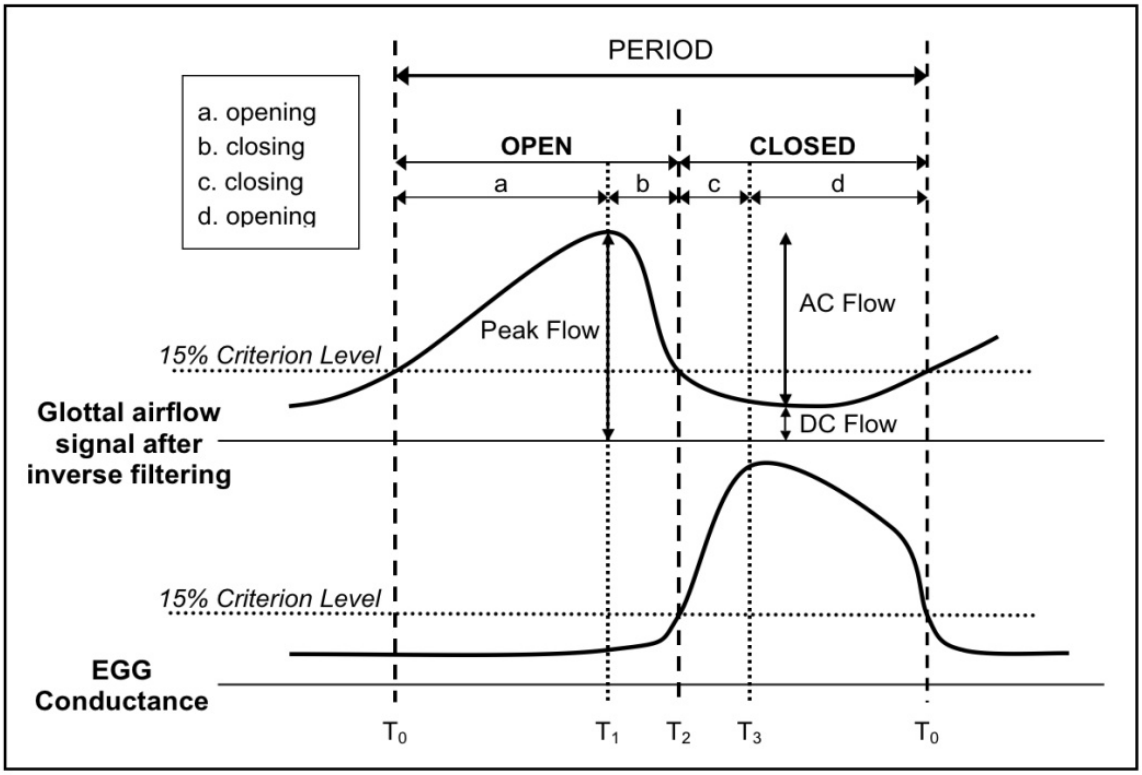

Figure 1. Part a, b, c, d of airflow and EGG signal along with visual representations of amplitude measures.

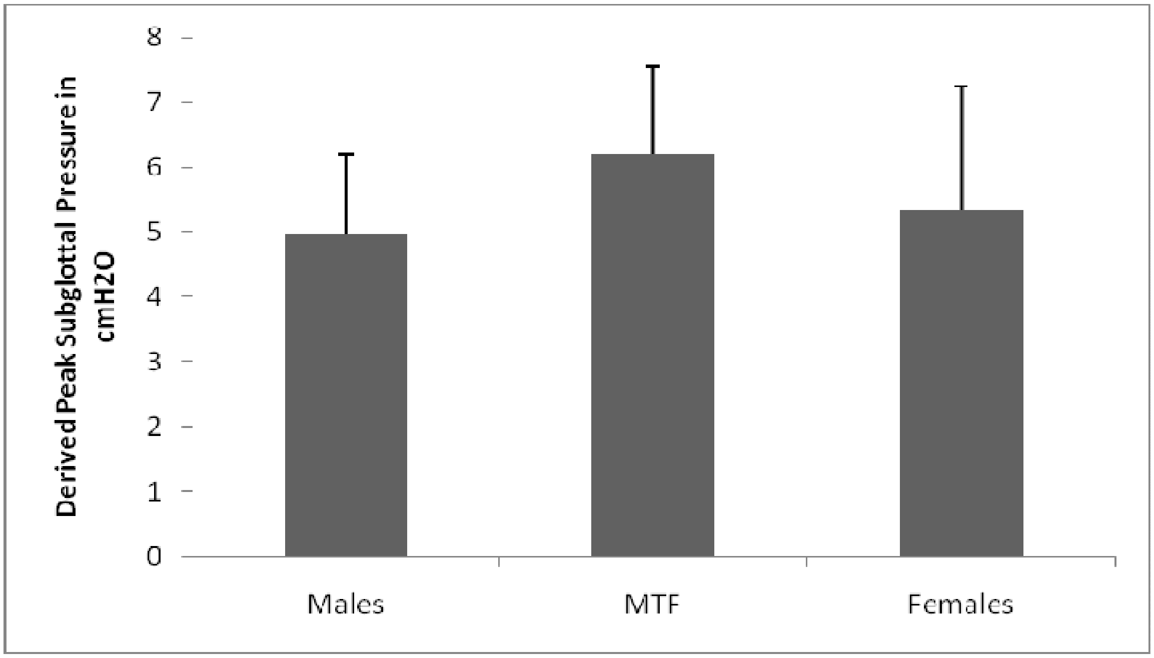

Figure 2. Mean derived subglottal pressure (+SD) in males (*n = 9*), females (*n = 9*), and male to female (*n = 10*) speakers.

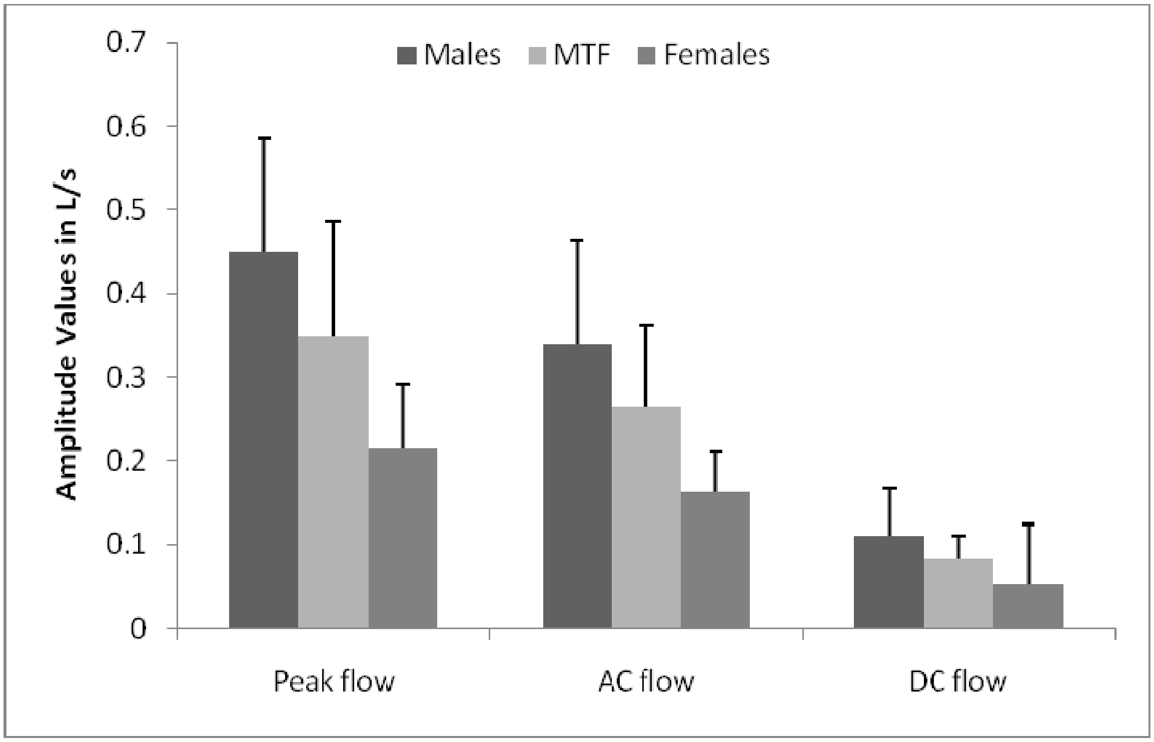

Figure 3. Mean airflow amplitude measures (+SD) for peak flow, DC flow, and AC flow for males (*n = 10*), females (*n = 10*), and male to female (*n = 10*) speakers.

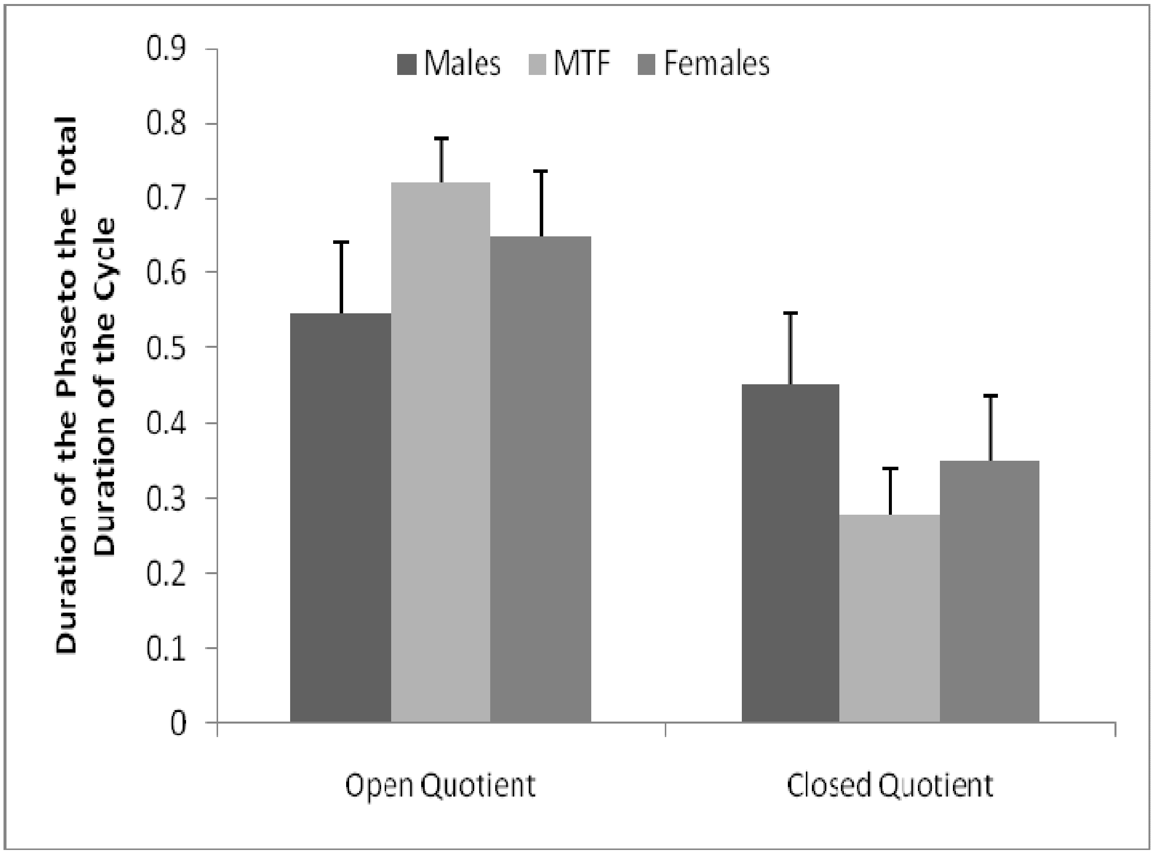

Figure 4. Graph of mean airflow duration quotient values (+SD) for open quotient and closed quotient for males (*n = 10*), females (*n = 10*) and male to female (*n = 10*) speakers.

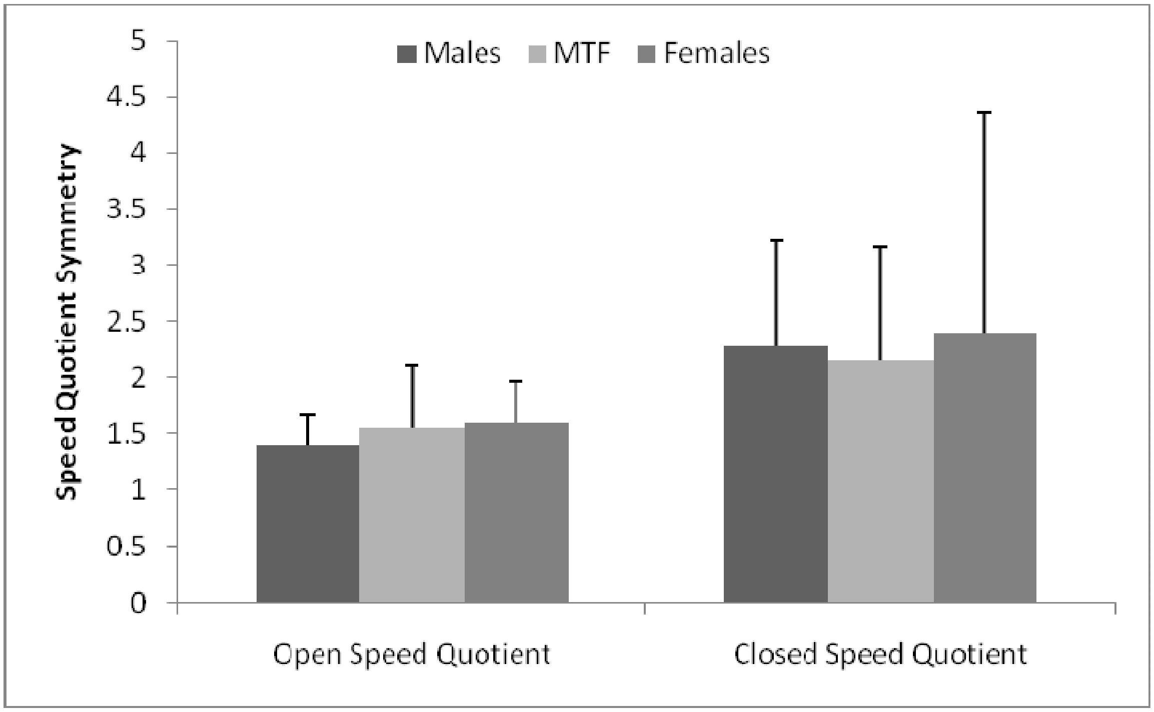

Figure 5. Graph of mean speed quotient values (+SD) for open speed quotient and closed speed quotient for males (*n = 10*), females (*n = 10*), and male to female (*n = 10*) speakers.

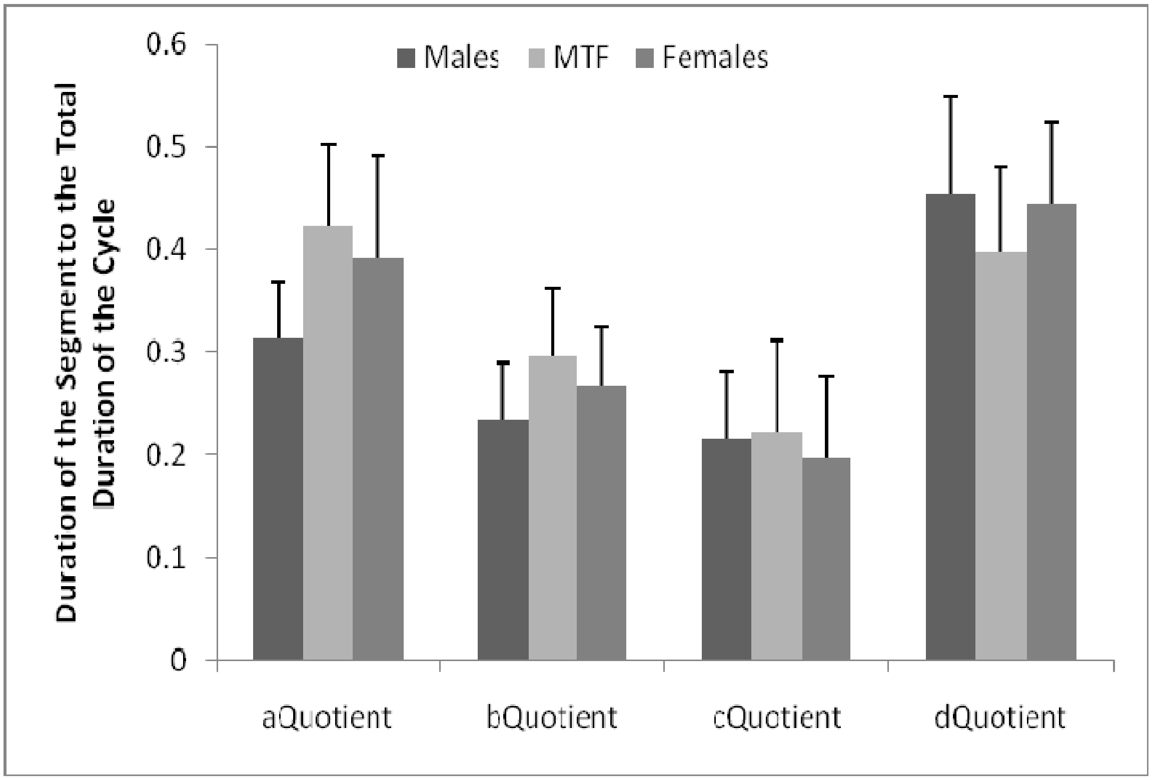

Figure 6. Graph of mean quarter quotient values for aQuotient, bQuotient, cQuotient, and dQuotient (+SD) for males (*n = 10*), females (*n = 10*), and male to female (*n = 10*) speakers.

Table 1

Airflow Amplitude Measures (*Perkell, Hillman, Holmberg & Gress, 1994*)

Measure	**Definition**	**Formula**	**Male (M)**	**Female (F)**	**Gender comparison**
Peak flow (L/s)	Peak flow is the greatest airflow at one time.	YT1 – zero	.410 L/s SD .087	.286 SD .1	M>F
AC (L/s)	AC flow is the difference between peak flow and DC flow in L/s.	(YT3-zero)/(YT1-zero)	.331 L/s SD .069	.156 SD .051	M>F
DC (L/s)	DC flow is the minimum airflow.	YT1-zero	.080 L/s SD .045	.130 SD .063	M<F

Table 2

Airflow Duration Measures (*Higgins & Saxman, 1991*)

Measure	**Definition**	**Formula**	**Male (M)**	**Female (F)**	**Gender comparison**
Open quotient	The time that the vocal folds are open over the duration of the cycle including both the opening and closing portions of the open segment.	(a+b)/period	57 SD 10.7	63 SD 9.7	M<F
Closed quotient	The time that the vocal folds are closed over the duration of the cycle including both the closing and opening portions of the closing segment.	(c+d)/period	50.8 SD 9.1	44.5 SD 11.6	M>F
aQuotient	This is the duration of the rise of the open segment on the airflow waveform, or a, over the period of the cycle.	a/period	No data available	No data available	No data available

bQuotient	This is the duration of the fall of the open segment on the airflow waveform, or b, over the period of the cycle.	b/period	19.7 SD 3.5	23.7 SD 3.2	M<F
Open phase Speed Quotient (OSQ %)	The time that the vocal folds are opening over the time that the vocal folds are closing during the open segment..	a/b	1.52 SD .35	1.36 SD .48	M>F

Table 3

EGG Duration Measures (*Chen, Robb, & Gilbert, 2002*)

Measure	**Definition**	**Formula**	**Male (M)**	**Female (F)**	**Gender comparison**
cQuotient	This is the duration of the rise of the closed segment on the EGG waveform, or c, over the period of the cycle	c/period	4.81 ms SD 1.46	1.86 ms SD .53	M>F
dQuotient	This is the duration of the fall of the closed segment on the EGG waveform, or d, over the period of the cycle	d/period	6.44 ms SD 1.46	3.46 ms SD 0.57	M<F (for quotient)
Closed phase speed quotient (CSQ)	The time that the vocal folds are opening over the time that the vocal folds are closing during the closed segment.	d/c	1.39 SD .27	2.09 SD .83	M<F

Note. Normative data from Chen, Robb, and Gilbert (2002) is measured in seconds and is therefore representative of the segment and not the quotient. This study will examine the quotient because the segment is highly dependent on fundamental frequency and quotient controls for this by dividing by the period of each cycle.

Table 4

Demographic Characteristics of Participants

	Age	Height	Weight	Birth gender	Desired gender	Medication	Prior voice treatment
MTF001	39	73	239	male	female	Estrogen, Estradiol, Spironlactone, Prometrium, Finasteride	No
MTF002	42	74	202	male	female	Estrogen, Keppra, Lamictal, Trileptal	Yes
MTF005	53	67	165	male	female	Lupron, Delestrogen, Paxil	Yes
MTF006	28	73	165	male	female	Estradiol, Aspirin, Fishoils, Vitamin C, Multivitamin	Yes
MTF008	59	67	155	male	female	Evamist	Yes
MTF009	50	68	200	male	female	Estradiol,	Yes

						Multivitamin, Antihistamine	
MTF010	57	69	168	male	female	Estradiol, Prilosec	yes
MTF011	55	74	238	male	female	Spironolactone, Premarin, Baby Aspirin	Yes
MTF012	24	71	160	male	female	Estradiol, Celexa, Wellbutrin	Yes
MTF014	29	69	160	male	female	Claritin, Estrace, Spironolactone	Yes
M001	28	71	145	male			
M002	53	70	220	male		Vitoran	
M004	28	74	210	male			
M006	43	71	157	male		Finasteride	
M007	53	73	245	male			
M008	22	71	185	male			
M009	39	68	140	male		Finasteride	
M010	59	69	165	male			
M011	56	71	190	male		Crestor	
M012	54	68	240	male		Meripax,	

					Aspirin, Lexapro
F003	28	64	135	female	Ortho tri cyclene, omeprazole
F004	24	64	135	female	Yasmine
F005	59	64	139	female	Synthroid
F006	27	65	155	female	Yasmine
F007	56	66.5	149	female	Prozac
F008	60	68	207	female	Atenolol
F009	37	70	170	female	
F010	40	64	205	female	Claritin
F011	59	66	183	female	Advair, Atrovent, Metformin, Tricort
F012	61	66	150	female	Synthroid

Table 5

Mean and Standard Deviation for Derived Subglottal Pressure, Airflow Amplitude Measures, and Open and Closed Quotients for Males, Females, and Male to Female Speakers

	Subglottal pressure (cmH_2O)	AC flow (L/s)	DC flow (L/s)	Peak flow (L/s)	Open quotient	Closed quotient
Male	4.968 (1.225)	.339 (.124)	.111 (.056)	.449 (.136)	.547 (.096)	.453 (.096)
Female	5.332 (1.923)	.162 (.048)	.052 (.072)	.215 (.077)	.649 (.086)	.350 (.086)
MTF	6.195 (1.362)	.264 (.087)	.083 (.08)	.347 (.141)	.721 (.061)	.279 (.061)

Table 6

Mean and Standard Deviation for Speed Quotients and Quarter Quotients for Males, Females, and Male to Female Speakers

	Open speed quotient	Closed speed quotient	aQuotient	bQuotient	cQuotient	dQuotient
Male	1.395 (.270)	2.273 (.954)	.313 (.055)	.234 (.055)	.217 (.065)	.454 (.095)
Female	1.601 (.372)	2.396 (1.971)	.392 (.099)	.267 (.059)	.196 (.080)	.445 (.078)
MTF	1.548 (.556)	2.149 (1.022)	.424 (.079)	.297 (.065)	.221 (.091)	.399 (.081)

Table 7

ANOVA Results for all Measures

Source	*df* (group, error)	F	*p*
Subglottal pressure	2, 25	1.634	0.215
Peak flow	2, 27	9.356	0.001
AC flow	2, 27	9.302	0.001
DC flow	2, 27	1.742	0.194
Open quotient	2, 27	11.288	0
Closed quotient	2, 27	11.288	0
Opening speed quotient	2, 27	0.659	0.526
Closing speed quotient	2, 27	0.078	0.925
aQuotient	2, 27	5.118	0.013
bQuotient	2, 27	2.764	0.081
cQuotient	2, 27	0.293	0.749
dQuotient	2, 27	1.192	0.319

Table 8

The p-values for Between-Group Differences for Derived Subglottal Pressure, Airflow Amplitude Measures, and Open and Closed Quotients for Males, Females, and Male to Female Speakers

	Subglottal pressure	Peak flow	AC flow	DC flow	Open quotient	Closed quotient
Male and female	1	**.001**	**.001**	.219	**.028**	**.028**
Male and MTF	.277	.214	.243	1.000	**.000**	**.000**
Female and MTF	.691	.065	.059	1.000	.193	.193

Table 9

The p-values for Between-Group Differences for Speed Quotients and Quarter Quotients for Males, Females, and Male to Female Speakers

	Open speed quotient	Closed speed quotient	aQuotient	bQuotient	cQuotient	dQuotient
Male and female	.836	1	.103	.713	1	1
Male and MTF	1	1	**.013**	.079	1	.487
Female and MTF	1	1	1.000	.789	1	.712