Voice Quality And Gender Identification: Acoustic And Perceptual Analysis

by

Courtney Cain Porter

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
June 2012

DALHOUSIE UNIVERSITY

SCHOOL OF HUMAN COMMUNICATION DISORDERS

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled “Voice Quality And Gender Identification: Acoustic And Perceptual Analysis” by Courtney Cain Porter in partial fulfilment of the requirements for the degree of Master of Science.

Dated: June 19, 2012

Supervisor: ____________________________________

Readers: ____________________________________

____

DALHOUSIE UNIVERSITY

DATE: June 19, 2012

AUTHOR: Courtney Cain Porter

TITLE: Voice Quality And Gender Identification: Acoustic and Perceptual Analysis

DEPARTMENT OR SCHOOL: School of Human Communication Disorders

DEGREE: MSc CONVOCATION: October YEAR: 2012

Permission is herewith granted to Dalhousie University to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions. I understand that my thesis will be electronically available to the public.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

The author attests that permission has been obtained for the use of any copyrighted material appearing in the thesis (other than the brief excerpts requiring only proper acknowledgement in scholarly writing), and that all such use is clearly acknowledged.

____

Signature of Author

DEDICATION PAGE

For my Mom and Dad, who taught me how to work hard and that, everyone deserves their own voice.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The voice is a fundamental method of communication and as such, helps in our efforts to define our identity. Projection of the appropriate voice is crucially important to transgender individuals in transition for acceptance as their identified gender. This study attempts to identify and examine the relationship between acoustic measurements of voice quality and the perception of speaker gender from audio recordings, including the male-to-female transgender voice, based on several acoustic properties that have been identified by previous studies. Recordings of female, male and transgender voices were acoustically analyzed for properties relating to differences in voice quality between men and women. Listeners then identified the gender of the recorded voices, with the intention of evaluating which voices are perceived as either male or female along with a corresponding rating of masculinity or femininity. *What acoustic measurements of voice quality cue listeners to gender and do they correlate with gender perception?*

LIST OF ABBREVIATIONS USED

f_0	fundamental frequency
H1	amplitude of the first harmonic
H2	amplitude of the second harmonic
F_1	first formant
F_2	second formant
F_3	third formant
BW	bandwidth
A1	amplitude of the first formant
A3	amplitude of the third formant
CPP	cepstral peak prominence
CPPs	cepstral peak prominence smoothed
SNR	signal-to-noise ratio
HNR	harmonic-to-noise ratio
SPI	soft phonation index
VTI	voice turbulence index
MtF	male-to-female

ACKOWLEDGEMENTS

I would love to thank all those that have supported me with their time and energy over the past two years of this project.

I am very grateful to my thesis supervisor, Dr. Michael Kiefte, who has shared his knowledge and experience as well as provided me with support and encouragement throughout.

Thank you to all my participants, rater and voice subjects, for your contributions and the time you donated to this work. I would also like to acknowledge Dr. Steve Aiken for the generous contribution of his hearing lab for the second half of the project. Thank you to the Youth Project for supplying lots of useful knowledge about the transgender community and encouragement.

To my supervising committee members, Dr. Joy Armson and Heather MacLean, and to Glen Nowell, thank you for your suggestions, advise, and support. A special thanks to Heather for helping me connect to the transgender community.

And to my friends and family for all your patience and support: believing I could achieve this goal even before I knew how.

For my husband, Joel, you are the most supportive and patient person to walk this earth, as well as the only man that could make stats seem fun.

CHAPTER 1: INTRODUCTION

The present study aims to contribute to a better understanding of the relationship between the perception of gender and voice quality, not only through perceptual means but through objective measurements as well. What acoustic cues the listener associates to one gender is based on an idea constructed by the general population. Where sex refers to the biological and physiological characteristics that define men and women, gender is a socially constructed role defined by a combination of behaviours, attributes, and activities society associates as appropriate to men and women (WHO, 2012). Through a complex interaction and combination of these characteristics gender perception occurs.

Research in voice and gender has shown that the male and female voice differ in a variety of ways. A population for which these differences are particularly important is the transgender community. Transgenderism is a complex condition involving a person's complete identification with the opposite sex of their birth (McNiell, Wilson, Clark, & Deakin, 2008). Although voice development is relevant for the female-to-male transgender community, often hormone supplements are effective in adjusting the voice to be perceived as male. As such, the current study will focus on solely the male-to-female (MtF) transgender voice. Studies involving the male-to-female transgender community have examined the association of different speaking and voice parameters – i.e. frequency, formants, intonation, voice quality, along with other pragmatic and paralinguistic components of speech - to the perception of the female gender (Adler, 2006). Although certain voice qualities, e.g. breathiness, have been implicated as related to the female voice and perception of female gender, few have systemically examined the relationship between the perception of femininity of the voice and its relationship to specific acoustic properties associated with vocal quality.

Previous studies have investigated the perception of vocal parameters associated with men and women. Fundamental frequency and its psychoacoustic correlate pitch, has been implicated as a significant indicator of gender in numerous investigations (Andrews & Schmidt, 1996; Coleman, 1983; 1975; Gelfer and Schofield, 2000; Spencer, 1988). Specifically, the female voice has been implicated as having a higher speaking pitch than the male voice. Nonetheless, research has indicated other vocal parameters, including perceived voice quality, influence gender perception for men and women. Consideration of how accepted parameters of voice quality relate to specific acoustic properties, how they are perceived, and how they can be measured is important to this investigation. A wide variety of acoustic measurements have been found to be associated to voice qualities and the definition of certain voice qualities (e.g. breathiness, hoarseness) remains ill-defined. The perception of many voice qualities may be as complicated as the perception of femininity itself. Consequently, a variety of objective measurements will be included in the current study.

A review of the research regarding male-to-female transgenderism and acoustic perceptions related to gender will be discussed, followed by a description of the method of the current study. Finally, the results of the investigation will be reported and discussed. Observations made through the process of this study will help in guiding the approach to transgender voice therapy, and the hierarchy of vocal characteristics included in therapies developing a female voice.

CHAPTER 2: REVIEW OF THE LITERATURE

2.1 TRANSGENDER POPULATION

The perception of gender identity based on the speaking voice is rarely a concern for the general population. Not many people wonder what components of his or her voice achieve the desired expression of gender. Whereas most of us may be frustrated by certain stereotypes and limitations attributed to our sex, we do not necessarily need to volitionally demonstrate our gender to the public. For those who are transgendered, a marked effort must be made to be accepted as the desired or identified gender (McNiell, Wilson, Clark, & Deakin, 2008).

Currently, the transgender community is included in two diagnostic categories in the proposed criteria of the 5th edition of Diagnostics and Statistical Manual of Mental Disorders (DSM-V): Gender Dysphoria (formerly Gender Identity Disorder or GID) and Transvestic Disorder (formerly Transvestic Fetishism; Winters, 2011). Where as a diagnosis of transvestic disorder primarily involves sexuality, a diagnosis of a gender dysphoria may not include the person's sexuality but is a marked incongruence of at least 6 months duration between one's expressed gender and biological gender as manifested by certain indicators (i.e., a strong desire to be the other gender or to possess primary or secondary sex characteristics of the other gender; APA, 2011). A diagnosis of GID requires that the individual have:

- strong and persistent cross-gender identification
- a persistent discomfort with his or her sex

- the disturbance is not concurrent with a physical intersex condition
- the disturbance causes clinically significant distress...(Vance, 2010).

Transgender individuals can experience an emotionally painful existence as well as anxiety and confusion if an endeavour to 'transition' to the identified gender is not attempted (Gender.org, 2001). The process of transitioning for a transgender individual is to attempt to attain congruency between the gender they identify with and the gender they are accepted as publicly (Dacakis, 2002). The transgender person will change their appearance, their name, and often undergo hormone treatment and/or sex reassignment surgery.

Unlike female-to-male transgender individuals, the male-to-female (MtF) transgender community will not experience a significant voice change due to the hormone therapy (Gunzburger, 1995). To develop a voice appropriate to their gender, MtF transgender individuals must seek to purposefully alter their speaking voices through therapy, surgery, or on their own. In certain situations where visual aid is unavailable to the listener, such as on the telephone, MtF transgender individuals can be mistaken for the inappropriate gender. Hancock, Krissinger, and Owen (2011) concluded from their investigation of quality of life of the MtF transgender person that their quality of life is related to both the speaker's self-perception of their voice and how others perceive her voice. Addressing components of the voice relevant to feminine perception is important not only in effective transition but also in the quality of life of a MtF transgender individual (Hancock, Krissinger, and Owen, 2011). The current study will focus on the male-to-female transgender community as the voice is not affected by hormone therapy and will more often need assistance in training to be perceived as the desired gender.

The role of voice therapy or training with the transgender client is to address the issue of gender misperceptions in public as well as in those situations without visual aid, such as in interactions over the telephone (Andrew & Schmidt, 1996). Both physical appearance and voice interact in a person's perception of a speaker's gender, such as if the voice heard cues the listener to one gender but the listener is presented with a visual aid incongruent with that gender, then the perception of femaleness will be unsuccessful (Van Borsel et al., 2001). As described by McNeill (2006), "... a less-than-acceptable physical appearance can be compensated for by attention to voice quality" (p.727). The aim of the speech-language pathologist in treating a transgender individual is to work towards maintaining a healthy voice within the appropriate frequency ranges and paralinguistic aspects acceptable to their gender identity (McNeill, 2006).

The voice is well established as not only one of the most difficult features to change in the transgender client, but one of the most crucial elements in gaining acceptance. Pitch, as measured objectively by fundamental frequency (f_0), is the primary focus of therapy with the transgender population (Van Borsel et al., 2001). Voice quality encompasses a wide range of possible meanings going well beyond vocal pitch. Aspects of voice quality include breathiness, nasality, hoarseness, and vocal fry. The term may be related to vocal fold vibratory characteristics, losses or gains of harmonic and inharmonic energy, as well as fluctuations in amplitude and frequency (Eskenazi, Childers & Hicks, 1990). Voice quality is a less tangible measure than f_0 but has been shown to be a distinct characteristic in the difference between what we perceive as a male or female voice.

An effective option for changing the speaking frequency of the MtF transgender individual includes surgery. These procedures have been found to be successful at increasing the mean speaking f_0 of MtF transgender individuals; an increase ranging from 16 to 131Hz. However, these procedures may not address the many characteristics of the voice that influence gender perception which complicate the successful alteration of the voice (Van Borsel et al., 2008).

There are a few types of surgeries available to transgender individuals that may increase their fundamental frequency (e.g. cricothyroid approximation; vocal-cord shortening; laser assisted voice adjustment; thyroid cartilage and vocal cord reduction); however, some voice feminization surgery has been found to be unacceptably aggressive, invasive, and inconsistent in its outcomes (Lawerence, 2004). Some surgeries can cause inconsistency between f_0 and resonances of the client's voice (Gunsburger, 1995). Gunzburger (1995) described one of the participants in her study of the transgender voice who had undergone surgical vocal cord construction as having "... an extremely high f_0 in the female mode (309Hz), which sounds meagre, unnatural, and falsetto-like." (p. 343). This indicates the importance of other vocal quality parameters, suggesting than pitch in this case is perhaps inconsistent with the other features of the voice characteristic of a gender. Improving the options available to transgender clients in a therapy setting is essential as better techniques of voice feminization surgery are still developing.

Van Borsel et al. (2008) concluded that, although voice feminization surgery—i.e., cricothyroid approximation—may raise the vocal pitch sufficiently; surgery alone may not be adequate in the creation of a voice that is perceived as totally female. Identification and exploration of these other voice characteristics to be modified will increase the likelihood the voice will not betray the transgender individual's biological sex. What these characteristics or possible gender cues of the voice are continue to be investigated in the literature.

2.2 VOICE QUALITY AND PARAMETERS IN GENDER

Vocal parameters, including pitch and quality, heard by a listener may be interpreted differently depending on the biological sex of the speaker (Murray & Singh, 1980). A man with a particular vocal quality may be interpreted differently – e.g. a feminine sounding male – than a woman with that same quality – e.g. a 'normal' feminine voice. The influence of biological sex on the perception of a speaker's voice is generalized to the entire male and female population in the Western culture – i.e. acoustic elements of the vocal parameters identifying speaker sex have been indicated as phoneme and speaker independent (Wu & Childers, 1991). Speaking fundamental frequency (f_0) as associated to the perception of pitch, formant frequencies (F), intonation, and voice quality (e.g., hoarse or breathiness) have been indicated as primary components of the perception of male or female speaker identity based solely on the acoustic signal (Andrews & Schmidt, 1996; Coleman, 1971, 1976; Gelfer & Schofield, 2000; Wolfe, Ratusnik, Smith, & Northrop, 1990; Wu & Childers, 1991).

Typically in voice training addressing gender perception, therapists try three methods: 1) raising the baseline f_0; 2) enhancing f_0 dynamics to produce an exaggerated intonation pattern by expanding the range of f_0; and 3) changing the baseline voice quality (Minematsu and Sakuraba, 2007). Although pitch has been identified as the foremost concern in altering the perceived gender of a voice, other vocal and non-vocal components of the speech also need to be addressed to present an accepted gender-appropriate voice (Van Borsel, Janssens & De Bodt, 2009; Adler, 2006; Mordaunt, 2006; Andrews & Schmidt, 1996). The perception of a speaker as either male or female is based on a gestalt comprised of many components beyond those previously mentioned above including: articulation, rate and intensity of speech, syntax, vocabulary, and pragmatics (Adler, 2006). However, components of pitch and voice quality are the main focus of exploration for the current study.

Fundamental frequency has a wide range of variation for both male and female speakers. The perception of habitual or average pitch perceived by a listener has been associated with the objective measure of speaking fundamental frequency (f_0). Adult males vary in habitual pitch between 107-146 Hz and adult females vary between 196-224 Hz (Colton, Casper, and Leonard, 2006). There is a wide gap between men and women for f_0 range referred to as gender-ambiguous pitch. This gap is important in the development of a female voice for the transgender client. Due to physiological constraints, a f_0 within the female range may not be accessible for a male transitioning to a female. A voice at 150Hz to 185Hz can still be perceived as female along with incorporating other aspects related to gender perception (Mordaunt, 2006).

Studies have indicated that the female voice is not merely a higher pitched version of the male voice. In a study of the acoustic and perceptual properties of male-to-female transgender individuals by Gelfer and Schofield (2000), only 2 of the 6 speakers with f_0 within a lower female range, as described by the authors, of 156-160Hz were identified as female. In another study of the transgender voice by Spencer (1988), only 2 of the 8 transgender individuals with f_0 of 160 Hz and higher were perceived as female. Spencer concluded that auxiliary cues which identified the transgender individuals as a male were available to the listener.

Physiological differences between the male and female vocal tract may be indicative factors as in the situations described above in the studies by Spencer (1988) and Gelfer and Schofield (2000). When f_0 is within the female pitch range and is still identified as male, formant frequencies may betray the biological sex of the speaker. Formant frequencies of the signal are dependent upon the anatomical differences of the vocal tract (Coleman 1971; 1976). Titze, as cited by Wu and Childers (1991, p.1828), indicated that the female vocal tract differed in length, thickness, and the resting angle of the glottis from the male vocal tract. Due to the relationship between formant frequencies and the vocal-tract size, (i.e. formants are inversely proportional to vocal-tract length) it is anticipated that formant frequencies would be a salient cue between male and female voices. The average female formant frequencies are 20% higher than those of male (Wu & Childers, 1991).

In a study by Pisanski and Rendall (2011) examining the relationship between f_0, formant frequencies, speaker size, masculinity and attractiveness, listeners consistently judged speakers with either low f_0 or low formant frequencies as larger and more masculine. The authors compared their findings for female and male voices and found that even female voices were indicated by listeners as more masculine and of larger size with low values of f_0 and formant frequencies. The authors suggested that these findings indicate that listeners use a general algorithm to assess body size and masculinity/femininity when listening to a voice, and not one specific to men or to women.

In a study by Perry, Ohde, and Ashmead (2001) on the acoustic bases of gender identification using children's voice, it was shown that perception of gender is dependent on several acoustic factors. Perry et al. (2001) indicated formant frequencies of the vocal tract were identified as acoustic properties with which listeners used to discern speaker sex of children when the f_0 was within a close range (e.g. children as young as 4 years old). This study was in agreement with several previous studies in which the fundamental frequency was factored out as a variable by using whispered, filtered, or electro-laryngeal speech (Coleman, 1971, 1976; Lass, Hughes, Bowyer, Waters, & Bourne, 1976). Other aspects involved in gender perception beyond f_0 and formants need to be examined – i.e. voice quality and intonation – because the perception of gender appears to be based on several acoustic factors (Perry et al., 2001).

The information a listener processes when judging the gender of a speaker may not simply be from the mechanism of the vocal tract and habitual pitch alone but from the behaviour of the voice—i.e. intonation. Wolfe et al. (1990) examined the role of intonation in their study of transgender male-to-female, female, and male voices.

Conversational responses were recorded for each subject. Their findings indicated that a voice perceived as female typically changed pitch more frequently, as the voice had a higher percentage of upward intonations and downward shifts than those perceived as male. By contrast, masculine voices were described as more monotone. In another study of normal female and male voices, as well as voices of gender-matched individuals with Parkinson's disease, all intonation parameters as measured by pitch range in semitones, mean f_0, standard deviation, as well as highest and lowest f_0 were found to be significantly elevated in female participants as compared with the male group (Skodda, Visser, and Schlegel, 2009).

Other studies disputed the claim that male and female voices differ in intonation style. Gelfer and Schofield (2000) found in their investigation of the perception of male-to-female transgender voices that intonation was not indicated as a significant factor of gender identification. In this study, the authors hypothesized that intonation was not significant because samples were not of spontaneous speech, but from readings of a passage. Content of speech and cultural background of the speaker may be factors influencing the perceived importance of intonation in gender identification. Primarily, English language intonation is thought to be constant across gender (Mordaunt, 2006).

The components and behaviour of fundamental frequency are the primary elements of gender perception; nevertheless, as stated by Mendoza, Valencia, Munoz, and Trujillo (1996) "...the differential synthesis of male and female voices implies much more than a mere scale of fundamental frequency" (p.61). The perception of voice quality has been found to depend on the speaker's sex as well as the type of qualities –

i.e. hoarseness, breathiness, nasality - of the voice (Eskenazi, Childers, & Hicks, 1990). Quality of the voice, such as breathiness or hoarseness, may be perceived as a normal characteristic, despite exceeding normative values, depending on the gender of the person speaking. Definitive definition and objective measurements of voice qualities are yet to be established and continue to be explored in the literature.

In a study by Eskenazi et al. (1990), acoustic measures of the speech signal were correlated with perceptual measures of voice quality including hoarseness, breathiness, roughness and vocal fry. For their investigation of voice quality, of the 6 measurements included in the study four were found be associated with voice quality: first, the spectral flatness of the residue signal (SFR), which measures the masking of f_0 harmonics by noise; pitch amplitude (PA), which measures the maximum amplitude of the normalized autocorrelation function of the residue signal and represents the degree of voicing in the signal; harmonic-to-noise ratio (HNR) which is a ratio of acoustic energy of harmonics to inharmonic noise; and percentage of jitter a measure of period-to-period variability of pitch. Their results showed that the perception of a hoarse voice was associated with a low PA and a high percentage of jitter; a breathy voice was characterized by a high percentage of jitter; a rough voice was correlated with a low spectral flatness of the residue signal and a low HNR; and finally vocal fry was indicated by a low PA and a low HNR. Eskenazi et al. (1990) did not, however, investigate the relationship between speaker sex and voice quality.

In a study by Awan and Roy (2005) of voice quality, breathiness and vocal roughness for women with functional dysphonia were frequently categorized as normal voice types by listeners. Rough voice was defined as ‘irregular vocal fold vibration’ and breathy voice as ‘turbulent noise originating from the glottis’. Acoustic measurements of shimmer and pitch sigma, the f_0 standard deviation converted to semitone range, were both correlated with a classification of rough voice. Breathiness was correlated with a significant increase in higher frequency aspiration noise content of the voice signal resulting in a steeper spectral tilt. The high frequency noise content is relatively weak in amplitude and possibly inaudible to the listener if not for the vibratory contribution of the vocal folds. The breathy voice quality may result in a spectrum with weaker high-frequency harmonics being replaced by high frequency aspiration noise (Klatt and Klatt, 1990). Spectral tilt or slope, as described by Awan and Roy (2005) is “...the relative spectral slope dependent on the degree of energy concentrated in the low- vs. high-frequency areas of the spectrum.” (p.277).

None of the voices classified as hoarse in the experiment were identified by listeners as a normal voice type for women’s voices. Awan and Roy (2005) found acoustic correlates for hoarseness included both period-based measurements (i.e. shimmer) and spectral-based measurements (i.e. cepstral peak prominence). Their study did not examine speaker gender identification specifically, though the authors suggested that a study involving male participants would be necessary to thoroughly investigate acoustic parameters predicting voice quality. The authors did note that a breathy voice is commonly misclassified as a normal female voice as opposed to a disordered female voice, especially when the voice signal has a relatively strong periodicity despite the additive noise of the turbulent airflow.

Acquiring more acoustic and perceptual information related to voice quality may help determine the vocal parameters a listener attributes to one gender versus the other. As well, further research may indicate the extent to which listeners accept specific voice qualities as typical of a gender, influencing the perception of a disordered voice - i.e. breathiness or hoarseness - from a typical voice. Finally, as in the case of the current study, further research may increase the understanding of the role voice quality has in speaker gender identification. Furthermore, finding the acoustic correlates to perceptual measures of voice quality will inform objective judgment of how effective or “passable” a voice is as one gender over the other.

Singh and Murray (1978) found in their examination of vocal quality in male and female voices that hoarseness was not a salient characteristic of a ‘normal’ female voice but one of the significant parameters of a ‘normal’ male voice. Hoarseness and the other qualities examined in this study by Singh and Murray were defined through the collective agreement of a group of speech-language pathologists chosen to rate each voice and its quality. In a follow-up study, Murray and Singh (1980) found a relationship between perceived nasality and effort of the voice and the perception of femininity.

Klatt and Klatt (1990) indicated in their study of the voice quality that a higher fundamental component, as measured as an increase in the amplitude of the first harmonic (H1), without aspiration noise may result in the perception of nasality. This is in contrast to perception of a breathy voice, which may be related to an increase in the amplitude of the first harmonic with aspiration noise. The first harmonic (H1) is the lowest frequency of the glottal source spectrum and is the same value as f_0 (Hixon, Weismer, and Holt, 2008). It may be possible that a vocal signal absent of high frequency noise (e.g. less aspiration noise) may be perceived as nasal. The perception of this particular voice quality, nasality, was possibly dependent on the actual speaker sex; however, in Klatt and Klatt's study nasality was not related back to the perception of gender.

Conversely, breathiness was identified in Klatt and Klatt (1990) and extensively in the literature as an important aspect of voice quality which may help listeners differentiate between male and female voices (Andrews & Schmidt, 1996; Hanson & Chuang, 1999; Hanson, 1997; Mendoza et al., 1995; Singh & Murray, 1978; Van Borsel et al., 2009). In research using aerodynamic analysis of the voice, higher airflow rates during phonation in female voices may be linked to a posterior glottal gap frequently exhibited by women, and is thought to be associated with a breathier voice quality (Gorham-Rowan and Morris, 2006). Breathiness is explained by Van Borsel et al. (2009) as "...the voice quality characterized by audible friction noise as a result of an incomplete closure of the vocal folds." (p.291). Dacakis (2002) describes breathiness, along with fundamental frequency, as having "...statistically significant correlations between perceptions of femininity in voice..." (p. 176). Both authors go on to express the importance of further investigation into the role of breathiness in speaker gender identification.

Klatt and Klatt (1990) studied voice quality variations across female and male voices by examining perceptual and acoustic measurements of breathiness and the potential correlation between breathiness to gender identity. In general, females were judged on average to be slightly more breathy than males. The authors investigated the acoustic correlates of breathiness among nine parameters, of which only two were found to be significant indicators of this voice quality. The results indicated that males and females differed in two perceptually important acoustic measures of breathiness: amount of aspiration noise in the third formant (F_3) region and in the relative amplitude of the first harmonic (H1). The amount of aspiration noise in the area of the third formant was related to the measure of spectral tilt in the signal, and the relative amplitude of the H1 was associated with the closure pattern (e.g. open quotient) of the vocal folds.

The opening and closing pattern of the vocal folds creates the shape of the glottal spectrum. The velocity at which the vocal folds return to the midline is related to the steepness of the glottal spectrum – i.e. decreasing amplitude or energy loss of the increasing harmonics in the spectrum. A longer open quotient is associated with a rapid reduction in energy across frequency or harmonics resulting in a steeper spectral tilt (Hixon, Weismer, and Holt, 2008). Indicated by Klatt and Klatt (1990), the most important factors correlated with the female speaking voice were a steeper spectral tilt and a longer open quotient of the vocal folds as compared to the male voice. Aspiration noise in relation to breathiness, along with the spectral tilt, was found to be important properties in synthesizing a natural, female voice. The authors concluded females are significantly breathier than males; however, a high degree of variation exists within the groups and in defining the term breathiness.

Hanson (1997) examined the acoustic parameters related to glottal characteristics associated with the voice quality of female speakers. The author described a theoretical model of breathiness in the female voice, based on previous reports of physiological and airflow data. Hanson predicted that female speakers were likely to have an incomplete closure of the vocal folds increasing the breathy quality of their voice and resulting in certain outcomes to the acoustic signal. The results indicated a relationship between the author's theoretical predictions of breathiness in the voice, with the relative amplitude of the first harmonic (H1), measured by comparing the amplitude of H1 to the amplitude of the second harmonic (H2), as well as an increase in bandwidths of the formants of the vocal tract. Larger difference between H1 and H2 amplitudes indicated an increased loss of energy in the lower frequencies attributed to a longer open quotient at the glottis. An increase in the first formant bandwidth, as measured by the difference in amplitude between the first harmonic and first formant (H1-A1), indicated an increase of noise in the signal also associated with breathiness.

One measurement related to voice quality and found to differ between men and women in the literature is the first formant (F_1) bandwidth (Hanson, 1997). Hanson found in her investigation of normal female voices that an association existed between the theoretical predictions of breathiness in the voice and increase bandwidths of the formants in the vocal tract. The author indicated that F_1 bandwidth may be associated with the pattern of glottal closure, specifically an incomplete closure pattern during a cycle of vibration. The greater the bandwidth of the formant the more noise in and dampening of the signal, which may be due to aspiration or breath in the voice (Hixon, Weismer, and Holt, 2008). A larger bandwidth at F_1 would indicate noise in the mid frequencies of the signal. When Hanson and Chuang (1999) examined the male voice, they found that F_1 bandwidth measurements were greater in the female voice than in the male voice.

In a follow-up study, Hanson and Chuang (1999) compared the glottal characteristics of the male speaking voice to the previous study on the female speaking voice. Female voices were found to differ in acoustic measurements compared with male voices in the following: relatively higher values of the first harmonic relative to the amplitude of the third formant (H1-A3; indicating a steeper spectral tilt), relative higher difference in amplitudes between the first and second harmonics (H1-H2; indicating a longer open quotient) and a greater degree of high frequency noise. Hanson and Chuang's findings of acoustic properties indicating differences between male and female voices were consistent with the results reported by Klatt and Klatt (1990). Hanson and Chuang indicated that the gender differences found by using their model of speech production were likely related to glottal configuration and vocal tract losses. In agreement with Klatt and Klatt (1990), Hanson and Chuang stated that perceived voice quality is affected by spectral tilt and that "...spectral tilt may greatly contribute to gender differences we perceive in speech." (p. 1077).

Nittrouer, McGowan, Milenkovic, and Beehler (1990) as well as Mendoza et al. (1996) both investigated the differences in voice quality between genders, and both found results in agreement with Klatt and Klatt (1990). Mendoza et al. (1996) evaluated data based on long-term average spectrum (LTAS) of male and female voices. Results indicated differences in the distribution of energy between sexes; in particular, they observed a steeper overall spectral tilt in the spectra of women's voices due to a greater concentration of energy in the higher frequencies. The authors believed this energy to be a consequence of greater levels of aspiration noise near the third formant. Their results agreed with the findings of Klatt & Klatt (1990) that "...acoustic characteristics of female voices lead to a "breathier" quality than in male voices." (Mendoza et al., 1996, p.64).

Nittrouer et al. (1990) found overall that in their analysis of male and female voices, women's voices had a greater amount of aspiration noise or turbulent noise than men. In an analysis computed separately for women's and men's voices, the authors found a strong relationship between the first to second harmonic ratio (H1-H2), an indication of harmonic energy in the vicinity of the glottis, and the signal-to-noise ratios (SNR). However, the relationship was different for men and women: A negative correlation was indicated for women and a positive correlation for men. For women, this indicated an increase in the amount of noise, assumed to be aspiration noise, as the relative amplitude of the fundamental increases. As well, Nittrouer et al. found that jitter has a significant relationship in the identification of speaker gender, as less jitter was found in the female than male voice.

The acoustic measurement of jitter in the voice has been shown by other authors to have a significant correlation to the perception of breathiness in the voice. Eskenazi et al. (1990) examined the relationship between perceived voice qualities and several acoustic measurements. Their results indicated that the percentage of jitter measured in the voice could predict the perception of breathiness. Shrivastav and Sapienza (2003) found a high correlation between the perception of breathiness and percentage of jitter ($r = .863$) as well as with SNR ($r = -.829$), while a moderate correlation was found between the perception of breathiness and the individual measurements of H1-H2, H1-A1, H1-A3.

In a study by Hillenbrand, Cleveland, and Erickson (1994), acoustic measurements were taken from normal male and female voices. Cepstral peak prominence (CPP), was indicated as having a strong correlation ($r = -.92$) to perception of breathiness. In this study, as well as in a follow-up study by Hillenbrand and Houde (1996) with pathological voices, the authors found a strong negative correlation between CPP measurements and perceived breathiness. The findings reported by Hillenbrand et al. were in close agreement to those of Klatt and Klatt (1990). Although male voices simulating a breathy voice were rated breathier than women, the authors hypothesized that listeners may have perceptually compensated for greater breathiness in the female voice. Perception of breathiness in the female voice may be skewed if breathy quality has been assimilated into Western culture as a feminine characteristic. Therefore the perception of simulated breathiness in female voice would have less of an effect compared to males.

CPP is a power spectrum of a log power spectrum—i.e. a spectrum of a spectrum. Periodic signals that show energy at harmonically related frequencies on the spectrum will show a strong component on the cepstrum corresponding to the regularity of harmonic peaks. Therefore, a well-defined harmonic structure will show a prominent cepstral peak. The amplitude of cepstral peak reflects both the level of harmonic organization and the overall amplitude of the signal. A linear regression is used to fit a line to the cepstrum. Cepstral peak prominence (CPP) is the difference in amplitude (dB) between the cepstral peak and the best fit linear regression of the entire cepstrum. Whereas periodic signals will have a greater deviation due to the harmonic structure of the harmonic spectrum—i.e. peak prominence—aperiodic signals will have a smaller CPP.

Another parameter, CPP-smoothed (CPPs) applies an additional step of smoothing the individual cepstra before extracting the cepstral peak and calculating the peak prominence (Hillenbrand and Houde, 1996). For smoothing, additional steps included a smaller measuring window – i.e. 2ms opposed to 10ms calculating CPP – and a two-step smoothing process: several cepstra are averaged across time by replacing the unsmoothed cepstral frames with the average of some cepstral frames to the left and right of the current frame. This process is followed by calculating a running average of cepstral magnitude or gamnitude from across the cepstral domain. CPPs is closely related to CPP but is less variable due to the smoothing process. This process is described in more detail in Hillenbrand and Houde (1996). Perceived breathiness in the voice is associated with a signal with low harmonic to noise ratio or a low CPP or CPPs.

The measurements H1-H2, H1-A1, H1-A3, SNR, jitter, and CPP as acoustic measurements associated with breathiness were all investigated by Shrivastav and Sapienza (2003) in female patients diagnosed with a voice pathology causing breathiness. All measurements were found to be correlated to the perception of breathiness, with CPP having the strongest correlation ($r = -.872$), followed by percentage of jitter ($r = .863$) and SNR ($r = -.829$). Examining breathiness, by way of objective acoustic measure and perceptual evaluation as a voice quality in the transgender voice has yet to occur frequently in the current literature. In a listener experiment conducted by Van Borsel et al. (2009), breathy samples of female voices were judged as more feminine than those non-breathy female samples. No acoustic analysis of the samples was undertaken and all data for breathiness in the voice was based on perceptual information. The authors noted the lack of studies involving transgender voices and voice quality, citing Andrew and Schmidt (1997) as the sole contributor in this area. Van Borsel et al. suggested that more research is needed with biological males and transgender individuals to further examine the relationship between increased breathiness and perceived femininity of voice.

Andrew and Schmidt (1997) identified breathiness, high pitch, and animation or increased intonation of the voice as most associated with the perception of femininity in their study of 11 heterosexual crossdressers. Although Andrew and Schmidt perceptually analyzed the crossdressers' voices for breathiness, no acoustic analysis was attempted. In a study of speech characteristics of male-to-female transgender individuals by Spencer (1988), only the fundamental frequency of the voice was examined as a factor indicating the gender of the speaker. Gelfer and Schofield (2000) hypothesized that breathiness may have been a contributing factor to the perception of transgender voices as female but were unable to draw any conclusions as they did not examine breathiness specifically. Gorham-Rowan and Morris (2006) concluded that those transgender individuals in their study producing a voice with increased laryngeal tension, combined with incomplete vocal fold closure, and therefore greater airflow during phonation, were successfully perceived as relatively more feminine than those who did not.

Singh and Murray (1978) found that breathiness in the voice had a positive correlation with the number of pitch shifts in the female voice. The shorter duration of the vocalic portion of speech resulted in the perception of a high degree of breathiness. The greater number of times that the pitch of the voice was shifted up or down decreased the vocalic portion or voiced component of speech, and increased the perception of breathiness in the entire sample of connected speech. The authors indicated that a greater degree of pitch shifting, or perceptually a more animated voice, may be more culturally acceptable in American females than males. The breathier quality of female speech, as investigated by Mendoza et al. (1996) and Van Borsel et al. (2009), has also been implicated in the acoustic analysis of Spanish and Dutch women, respectively. Although, breathiness as a cue to speaker gender may be culturally dependent, it has been indicated to be important in the culture of the transgender population that is the subject of the present study. If breathiness is associated with perceived femininity, incorporation of breathiness into the assessment of a passable transgender voice may be important to the acceptance of the transgender individual as their identified gender into Western cultural society.

2.3 JUSTIFICATION FOR THE STUDY

The voice is a multifaceted tool to which a person's identity, including personality, culture and gender, is strongly linked. Addressing gender congruency in voice therapy is complicated, as stated by Coleman (1983), in that "... the gender characteristic most resistant to convincing change is the voice." (p.293). Investigating the multiple components involved in gender perception from the voice will contribute to the knowledge of developing a female voice.

The acoustic characteristics of speech which listeners focus on to identify the gender of the speaker may not be related to one component of the voice. A variety of psychoacoustic properties are used by listeners in making the dichotomous choice of speaker gender. As a result, altering a voice to be effectively identifiable to the general public as the intended gender of the speaker is very difficult (Avery & Liss, 1996). To create a truly passable and appropriate voice that will satisfy the transgender client, all possible avenues of communication therapy must be examined. Perceptual and acoustic information must be obtained to better understand the differences between the male and female voice, as well as the fundamental aspects of the transgender voice to focus on in therapy. The information collected will inform the following research questions:

- What acoustic measurements of voice quality cue listeners to the perception of one sex or the other?
- Do acoustic measurements of voice quality correlate with the perception of a speaker's masculinity or femininity?

CHAPTER 3: METHOD

3.1 RECORDING AND ACOUSTIC INFORMATION

3.1.1 Participants

Participants were recruited from the general public in the Halifax Regional Municipality of Nova Scotia. Biological females and males were recruited through means of publication (e.g. public notice) and email. For the samples of male-to-female (MtF) transgender voices, individuals were recruited from the Nova Scotia Hearing and Speech clinic at the Dickson centre, Halifax, and with the assistance of community outreach programs for transgender people (i.e. the Youth Project[1]).

The speakers consisted of 30 biological female, 27 biological male, and 4 male-to-female (MtF) transgender individuals all of whom were native speakers of North American English. One biological male participant was recovering from puberphonia and due to his unique vocal condition was included with the transgender voices. The group of transgender voices and the participant with puberphonia will be referred to as the transgender/puberphonic voice group (abbreviated as transgender or transgen).

All participants were literate adults. Their ages ranged from 19 to 80 years (38.8 mean age). Mean age was 45.8 years for women (SD = 16.4), 42.4 years (SD = 19.7) for

---

[1] The Youth Project is a non-profit charitable organization providing support and services to youth, 25 and under, around issues of sexual orientation and gender identity. Located in Halifax, NS, the Youth Project provides leadership opportunities and programs to lesbian, gay, bisexual, and transgender youth

men, and 28.4 years for transgender women (8.5 SD). A one-way ANOVA found no significant difference among the age of recording participants (p = .130).

Male-to-female transgender participants were either in transition or living full-time in their female gender identity. Two transgender participants had completed some voice therapy in the past and another individual was receiving ongoing therapy. The fourth transgender participant was living full-time as a female but had not sought, nor was interested in, seeking voice therapy. The transgender woman currently in voice therapy produced both a voice associated with her male identity and her female identity. Both voices were again recorded: the male included in the male voice samples and the female included in the transgender voice sample.

The fifth participant included in the transgender/puberphonic voice group was a male with puberphonia, which is described as the persistence of a high-pitched voice beyond the age at which voice change is expected to have occurred (Colton, Casper, and Leonard, 2006). Although he was not a transgender individual, due to the hypothesized impact f_0 would have on gender perception, he was included with the transgender female group. Despite a high-pitched voice, men with puberphonia have a voice that is rarely described as feminine (Colton, Casper, and Leonard, 2006). This individual had received voice therapy and was, at the time of recording, able to produce a lowered pitch in voice. He did, however, express that he was still able to speak spontaneously in his high pitch voice. Both voices were recorded: the voice which had developed through therapy was included in male recordings and the puberphonic voice with the transgender voices.

Speakers answered questions regarding history with speech, language, and voice services, smoking history, and any other pertinent information (e.g., vocal or oration training). As well, any personal observations of current voice quality at the time of the recording were noted (e.g., recovery from cold). Each participant was assigned a number and letter depending on their gender (e.g., F1, M1, and T1). Details of participants are listed below in Table 1.

Table 1: Background information on different speakers

Participant	Age	Gender	Smoking or Non	SLP services	Region of childhood origin	Other vocal info
F01	32	Female	Non	SLP: childhood speech impediment	Maritimes	
F02	60	Female	Non	SLP: None	Maritimes	choir singer
F03	58	Female	Non	SLP: childhood speech impediment	Maritimes	
F04	22	Female	Non	SLP: None	Ontario	Professionally trained singer
F05	30	Female	Non	SLP: None	Maritimes	Trained singer
F06	22	Female	Occasional	SLP: None	Ontario	
F07	52	Female	Occasional	SLP: None	Maritimes	
F08	49	Female	quit 15 years ago	SLP: None	Sydney	
F09	68	Female	Non	SLP: None	Ontario	
F10	68	Female	Non	SLP: None	Ontario	recovery from cold
F11	69	Female	Non	SLP: None	Alberta	
F12	36	Female	Non	SLP: None	Maritimes	
F13	62	Female	Non	SLP: None	Maritimes	Raspy voice
F14	19	Female	Non	SLP: breathy voice - seen pre-nodules (3 years ago)	Maritimes	
F15	52	Female	Non	SLP: None	Ontario	
F16	58	Female	Non	SLP: None	Dartmouth	
F17	42	Female	Non	SLP: None	Saskatchewan	
F18	34	Female	Non	SLP: None	Cape Breton	Voice (for orating) training
F19	69	Female	Non	SLP: None	Halifax	
F20	50	Female	Non	SLP: None	Vancouver	
F21	65	Female	Non	SLP: None	Halifax	
F22	23	Female	Non	SLP :None	Halifax	
F23	30	Female	Non	SLP: None	Halifax	Recovery from cold
F24	59	Female	Non	SLP: None	Cape Breton	
F25	59	Female	Non	SLP: None	Nova Scotia - French Canadian	
F26	31	Female	Non	SLP: None	Maritime/Ontario	
F27	22	Female	Non	SLP: None	Nova Scotia	

Participant	Age	Gender	Smoking or Non	SLP services	Region of childhood origin	Other vocal info
F28	49	Female	Non	SLP: Voice therapy for hourglass VF	Cape Breton	Professionally trained singer; GERD
F29	46	Female	Non	SLP: None	Ontario	
F30	37	Female	Non	SLP: None	Halifax	
M031	30	Male	Non	SLP: None	Maritimes	Singer
M032	20	Male	Non	SLP: None	Maritimes	Professionally trained singer
M033	21	Male	Occasion	SLP: None	Maritimes	Professionally trained singer
M034	80	Male	recently quit	SLP: None	Maritimes	
M035	22	Male	Occasion	SLP: None	Maritimes	Professionally trained singer
M036	72	Male	Non	SLP: None	Saskatchewan	Hoarseness from Asthma
M037	60	Male	quit 25 years ago	SLP: None	Moncton	French accent (slight)
M038	62	Male	Non	SLP: None	Montreal	Recovery from cold
M039	64	Male	Non	SLP: None	Dartmouth	
M40	69	Male	smokes pipe	SLP: None	Ontario	Voice and speaking training
M41	53	Male	Non	SLP: MVA 1982 - TBI	Manitoba; living Halifax	
M42	38	Male	Non	SLP: articulation	Portland, Oregon; living Halifax	
M43	24	Male	Non	SLP: None	Maritimes	
M45	65	Male	Non	SLP: None	American (9 yrs in Canada)	
M47	53	Male	Non	SLP: As child pre/post cleft palate	Cape Breton	
M48	59	Male	Non	SLP: None	Ottawa	
M49	42	Male	Non	SLP: None	Cape Breton	
M50	22	Male	Non	SLP: None	Nova Scotia	
M51	22	Male	Non	SLP: None	Nova Scotia	
M52	68	Male	Non	SLP: None	Childhood in N. Eastern USA; 40yrs Maritimes	
M53	45	Male	Non	SLP: None	Cape Breton; living in Ontario	
M54	44	Male	Non	SLP: None	Halifax	
M55	30	Male	Non	SLP: None	Halifax	
M56	26	Male	Non	SLP: None	Dartmouth	
M57	26	Male	Non	SLP: None	Halifax	
M58	26	Male	Non	SLP: None	b. Germany (first 5 years); Raised in Colorado; living 2 yrs in Canada	
T59/M44	19	Male	Non	SLP: Voice therapy for puberphonia since Dec.	Nova Scotia	Puberphonic voice Not transgender
T60/M16	22	MtF	Non	SLP: Transgender voice therapy 1-2 months (ongoing)	Halifax	In transition - not full-time as woman/allergies
T61	39	MtF	Non	SLP: Transgender voice therapy 2-3 months	New Brunswick	9 years living as woman
T62	27	MtF	Non	SLP: None	Nova Scotia	5 years living full-time as a woman
T63	35	MtF	Light smoker	SLP: Transgender voice therapy several months	Halifax now living in Vancouver	5 months full-time as woman

3.1.2 Procedure: Recordings

Male-to-female (MtF) transgender, male, and female voice samples were recorded in quiet areas which were accessible to the participants (e.g. homes, libraries, churches, and School of Human Communication Disorders), using a Shure head-mounted microphone and a Marantz Pro digital recorder. Participants were given a consent form to read and sign. Prior to the start of the task, fully informed consent was obtained from each participant. Anonymity for each participant was assured as only the subject's age, sex, and number of years living as a transitioned female (when relevant) was recorded. The participants were given the opportunity to practice reading the Rainbow Passage, 15 randomly presented carrier phrases (i.e., *Please say the word /hVd/,* again) and sustaining /ɒ/ vowel until they felt comfortable in recording their voice. All material read by the participants is presented in Appendix 1. Each individual was instructed to perform the requested stimulus at a comfortable, conversational level. MtF transgender participants were given an additional instruction to use their "best" female voice or voice they feel would enable them to pass in a non-visual situation (i.e. over the telephone). For the carrier phrases, participants were instructed to prolong and emphasize the highlighted word as marked on the written cue card (e.g. had). Finally the participant was instructed to sustain the vowel /ɒ/ for a period of 10 seconds as indicated by the researcher.

3.1.3 Procedure: Acoustic Analysis

Goldwave program was used to isolate the spoken material for acoustic analysis and the listener experiment (Goldwave Inc., 2001). Fifteen carrier phrases, as presented in appendix 1, each containing a vowel - i.e. /ɒ/, /ɑ/, /æ/, /o/, /u/, /i/, /ɛ/, /ʌ /, /ɪ/, /ʊ/, /ɚ/ - or a diphthong - /ɑɪ/, /e ɪ/, /oʏ/. /ɑu/ - in /hVd/ production were recorded. Four carrier phrases sampling the four corner vowels were chosen for further analysis: *Please say the word hodd again*, *Please say the word heed again, Please say the word who'd again,* and *Please say the word had again.*

Carrier phrases were isolated by using the trim function and segments containing extraneous material (e.g. spontaneous questions, examiner voice, or dead air) were deleted. The individual carrier phrases were transcribed using the *textgrid* function in the computer-based analysis software, PRAAT (Boersma and Weenink, 2001). The vowels /ɒ æ/ were isolated from the words *had* and *hodd* from two of the carrier phrases. The decision to analyze only the vowels /ɒ æ/ is based on the findings of Nittrouer et al. (1990) in their examination of acoustic properties of male and female voices that low vowels are a reliable measurement of spectral tilt across speakers. The evaluation of acoustic differences in speaker gender identification using low vowels was also used by Klatt and Klatt (/ɑ/; 1990), Hanson (/æ/; 1997), and Shrivastav and Sapienza (/ɑ/; 2003). The remaining recorded material was not analyzed but was preserved for future study.

The voice samples were then acoustically analyzed using the computer-based analysis software, PRAAT (Boersma and Weenink, 2001). For the isolated vowels taken from the context /hVd/, the vowel midpoint was chosen for analysis as measurements would be least affected by the adjacent consonants (Klatt and Klatt, 1990). Measurements were taken of each sample's mean fundamental frequency (f_0) and its standard deviation from which the first two harmonics were calculated - i.e. f_0 is first harmonic and second harmonic is a doubling of f_0 (Hixon, Weismer, and Holt, 2008). Along with these measurements, the first three formant frequencies (F_1, F_2, F_3) were also measured from the vowel centre of /ɒ æ/ as sampled from the words *had* and *hodd* from within the carrier phrase, as well as from the sustained vowel recorded. The bandwidth (BW) of F_1 was measured as well due to its relationship to the presence of breathiness in the vowel — i.e. the greater the first-formant bandwidth, the greater the amount of breathiness in the sound (Hixon, Weismer, and Holt, 2008). Following identification of mean f_0, first harmonic, second harmonic, F_1, F_2, and F_3, the long-term average spectrum of the acoustic signal was determined using PRAAT to calculate the nearby amplitudes of first harmonic, second harmonic, F_1, and F_3. The amplitudes of the first and second harmonics are labeled as H1 and H2, and the first and third formant amplitudes are abbreviated as A1 and A3.

Corrections as described by Hanson (1997) were applied to H1, H2, and A3 prior to calculating the differences between H1-H2, H1-A1, and H1-A3. The corrections, as described by Hanson (1997), were made to account for the “boosting” effect the first formant has on the vocal-tract transfer function, influencing the amplitude of the first and second harmonics; therefore correction, as calculated using the values of F_1 and f_0 is subtracted from the amplitudes of H1 and H2. As well, corrections are made to A3 to compensate for the effect of F_1, and F_2. The amplitude value of F_3 was increased by using a log equation involving the values of F_1, and F_2 from the specific vowel (e.g. /ɒ æ/) as well as the values of F_1, and F_2 taken from a neutral vowel. Formants taken from neutral vowels were estimated by averaging the values of F_1, and F_2 taken from the four corner vowel samples as recorded using the carrier phrases. These corrections would allow for comparison across speakers and vowels.

A computer system for voice analysis, the Computerized Speech Lab (CSL) 4500 by KayPENTAX Corp. (2003) was also used in analyzing the acoustic signal and to determine the semitone range and standard deviation. Specifically, the Multi-Dimensional Voice Program (2004), found in CSL, was used to analyze the percentage of jitter and shimmer, as well as the noise parameters of soft phonation index (SPI), voice turbulence index (VTI), and noise to harmonic ratio (NHR). These measurements were taken at the vowel in the /hVd/ from the carrier phrase recordings as well as in the sustained vowel.

The noise parameter measurements, SPI and VTI were taken to investigate the average ratios for: the low-frequency harmonic energy (70-1600 Hz) to high-frequency energy (1600-4500 Hz); and the spectral inharmonic high-frequency energy (2800-5800 Hz) to the spectral harmonic energy in the range 70-4500 Hz, respectively. The measurement of VTI examines the relative energy level of high-frequency noise and harmonics. A general measure of noise in the signal, NHR, records the average ratio of the inharmonic spectral energy in the frequency range 1500 to 4500 Hz to the harmonic spectral energy in the frequency range 70-4500 Hz.

Acoustic parameters previously found in the literature to discriminate male and female voices were also calculated (Klatt and Klatt, 1990; Mendoza et al., 1995; Hanson, 1997; Hanson and Chuang, 1999; Shivastav and Sapienza, 2003). This includes the relative amplitude of the first harmonic (H1) in /ɑ æ/ as determined by the difference in amplitude relation between H1 and H2. The amplitude relations involving the difference between the amplitude of the first harmonic (H1) to the amplitude of the second harmonic (H2), and the difference between H1 to the amplitude of the first formant (A1) were determined to estimate the duration of the open quotient of the glottis and the bandwidth of F_1, respectively. In addition, spectral tilt was estimated using the difference between the amplitudes of H1 and the third formant (A3) of /ɑ æ/ (Klatt and Klatt, 1990). All of these measures have been described as correlates to perceived breathiness.

Cepstral peak prominence (CPP) was calculated using the computer-based analysis software SpeechTool (Hillenbrand, 2008). The CPP and CPPs was determined for the entirety of each carrier phrase and sustained vowel production.

3.2 PERCEPTUAL EXPERIMENT

3.2.1 Participants

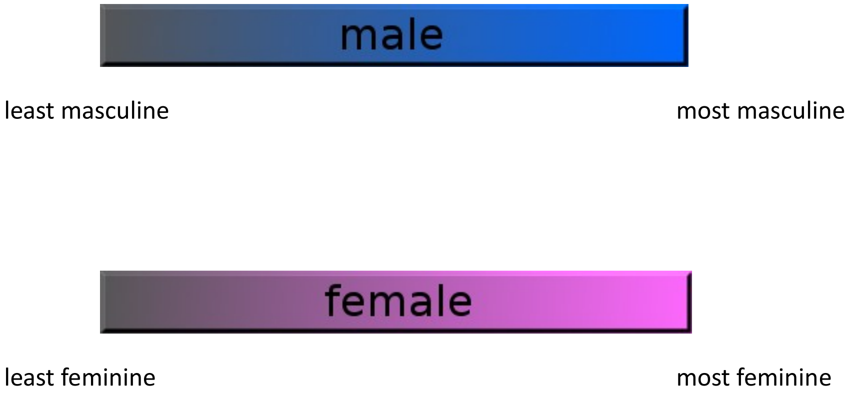
Twenty-six listeners, 23 females and 4 males, with a mean age of 26 years (range 22 to 33 years, SD = 2.94) participated in this experiment. None of these listeners had participated in the recording of part one. All participants, except one male, were recruited from the graduate program in Human Communication Disorders at Dalhousie University in Halifax, Nova Scotia. All listeners were literate adults who were native speakers of North American English. All participants were screened for normal hearing and all except one male were found to have average binaural hearing. This male participant had a slight threshold shift in his right ear due to impacted wax and normal range of hearing in his left ear. Since hearing was normal in one of the participant’s ears he was included in the study.

3.2.2 Procedure

The perceptual experiment was conducted at the School of Human Communication Disorders, Dalhousie University in the Speech-Perception Laboratory. Fully informed consent with a description of the task was provided before the participant had started. The listeners were told that transgender voices were used in the recordings following their ratings of the voices; however, there was no attempt to blind the participant from this fact prior to their participation. After informed consent was acquired the subject was screened for normal hearing.

Following the hearing screening and normal range of hearing was established, the examiner reinstructed the participant in the task. A sample screen, as depicted in figure 1, of the two scales for masculinity and femininity, graded blue and pink bars respectively, was shown to the participant. The participant was instructed to first make a judgement as to whether the voice they had heard was male or female, and then to indicate by clicking within one of the buttons along the corresponding scale, male or female, how masculine or feminine they judged the voice to be.

Figure 1: Sample screen of gender perception selection



Many studies have utilized scales of gender in the collection of transgender voice identification. Andrews and Schmidt (1996) used 18 perceptual rating scales with 8-point semantic differentiation of femininity-masculinity developed by Gelfer (1988). Gelfer and Schofield (2000) required the participants in their listening experiment to identify each speaker as female or male, and complete a 7-point rating scale of very feminine to very masculine. Spencer (1988) also used a dichotomous choice of sex as well as a scaled perception of sex, i.e. masculinity/femininity. For the current study the measure of masculinity/femininity was recorded across the bars from 0 (least masculine/feminine) to 100 (most masculine/feminine). Femininity scores were arbitrarily recoded as negative (0 least feminine to -100 most feminine) and masculine scores as positive (0 least masculine and 100 most masculine).

The listeners' judgements were recorded by MATLAB (Mathworks, Inc.). The listener was given control over the input using a laptop computer. TDH headphones were placed by the listener comfortably onto their head. Each sample was played consistently at a comfortable listening level. Upon entering the program each recording was delayed by 1 second following the listener's selection. The listening task was broken into three parts: the listener first judged segments of sustained vowels, followed by two carrier phrases. The listener was randomly presented with all recorded speakers saying first, *Please say the word had, again*, followed by all recorded participants saying, *Please say the word hodd, again*. Each speaker was heard by a listener once saying each of the carrier phrases. The listener heard 63 randomly presented voices saying each carrier phrase, for a total of 126 connected speech stimuli.

Each recording of a sustained vowel was segmented into three parts with cuts occurring at quarter intervals with the first segment centred at 25% of the duration of each recording, the second centred at 50% of the recording length, and finally the last segment was centred at 75% of the original recording. The length of each segment was approximately 300 to 400-ms. All shortened sustained vowel segments were played for the listener, resulting in 189 segments, with each recorded participant repeated three times randomly throughout the listening task.

The entire listening task took around 40 minutes to complete depending on the listener's speed of judgement within each task. Short breaks were given between the three listening tasks – i.e. the two carrier phrase conditions, and the shortened sustained vowel segments. The listener's rating for gender (male or female) and judgement of masculinity/femininity were automatically downloaded to Microsoft Excel.

3.3 EXPECTED OUTCOMES

It was hypothesized that high fundamental frequency (above 160Hz) is a primary indicator of speaker gender identity (Andrews & Schmidt, 1996; Coleman, 1983; 1975; Gelfer and Schofield, 2000; Spencer, 1988). Voices with steeper spectral tilt and therefore more aspiration noise at higher formants were also expected to be correlated with perceived femininity (Hanson and Chuang, 1999; Hanson, 1997; Klatt and Klatt, 1990; Mendoza et al., 1996; Nittrouer et al., 1990). In addition, those voices with a higher percentage of jitter and a larger NHR than other voices were predicted to be breathy and therefore perceived as female (Eskenazi et al., 1990; Nittrouer et al., 1990; Shrivastav and Sapienza, 2003).

Aspiration noise as related to breathiness was expected to be significantly different between the perceived female voice and the male voice. Correlation between the perception of speaker gender and the measurements associated with breathiness in the voice was expected to be observed. It was expected the fundamental frequency would have the strongest correlation.

CHAPTER 4: RESULTS

4.1 OVERALL DATA

The mean and standard deviations of all the measurements of voice quality and frequency were calculated for each gender group (i.e. female, male, and transgender) across the three different production contexts (i.e. sustained vowel and two carrier phrases). As well, the mean gender rating and percent correct identification of gender choice made by listeners to each speaker's voice were calculated to determine the perceived gender of each voice from each context.

Tables of raw data, mean and standard deviation are presented in Appendix 2 for each gender group and context. Tables of listener response information are presented in Appendix 3. Pearson correlations between the various measurements taken from the voice samples and the listener responses for gender rating are presented in Appendix 4.

4.2 ACOUSTIC MEASUREMENTS OF VOICE QUALITY

The analyzed acoustic signals were taken from three different productions by speakers: a sustained vowel /ɒ/ and two carrier phrases, *Please say the word /hVd/, again,* containing either the word *had* or *hodd*. Measurements related to glottal configuration (H1-H2, H1-A1, H1-A3, F_1 bandwidth), noise parameters (VTI, NHR, SPI), fundamental frequency (f_0), and perturbation of the signal (jitter and shimmer) as described in chapter 3, were recorded from the segmented vowels of the words *had* and *hodd* as well as the sustained vowel. Measurements of f_0 range in semitones and Hertz were taken only from the carrier phrases to investigate intonation in connected speech. Measurements of aperiodicity in the signal, CPP and CPPs, were taken across the entire phrases and the sustained vowel productions.

The sample size of the transgender group was heterogeneous and grossly unequal (N=5) to the male (N=28) and female (N=30) groups, which would seriously affect the validity of an analysis comparing the three groups (Portney and Watkins, 2009). To determine how successful members of the transgender group were at presenting as females, transgender/puberphonic voice group data will be discussed individually in reference to those vocal elements identified as significantly different between male and female groups.

To establish the differences between the acoustic measurements for the voice between the two sexes, the means of female and male voices were compared using independent *t*-tests for each measure. Numerous *t*-tests were used to analyze these vocal elements and as such interpretation and any generalization of the results should be done so with caution due to the elevated risk of Type I error. Across all contexts – i.e. sustained vowel, and the two carrier phrases – three measurements were found to differ significantly between men and women: mean fundamental frequency (f_0), first formant (F_1) bandwidth, and cepstral peak prominence (CPP). These measurements will be presented first, followed by those measurements found to only differ significantly between men and women for certain contexts.

As expected a significant difference of mean f_0 ($p <.000$) was observed for all contexts. As well, first formant (F_1) bandwidth was observed to differ significantly for all contexts: sustained vowel (t=3.75, $d.f.$=44.08, p <.001), /æ/ from had (t=2.83, $d.f.$=33.39. p<.008), and /ɒ/ from hodd (t=2.58, $d.f.$=53.412, p <.013). As with f_0, the value of F_1 bandwidth was found to be higher in females than males. The mean and standard deviation for the male and female groups are presented in Tables 2, 3, and 4 for each recorded context.

In the results from the transgender/puberphonic voice group, as seen in appendix 2 tables 11, 12 and 14, the individual participant with the highest f_0 across all contexts was participant T59. This participant also had the highest value of F_1 bandwidth for the measurement taken from isolated /ɒ/. The participants varied for whom F_1 bandwidth was a highest of the group in the remaining contexts: T63 had the highest value when measured from the sustained vowel and T60 for the F_1 bandwidth measured from /ɒ/. The measurements of frequency and F_1 bandwidth for all three groups are illustrated in the means Boxplot depicted in figure 2 and 3, respectively. Outliers for each group are labeled with the participant's number.

Figure 2: Boxplot of mean f_0

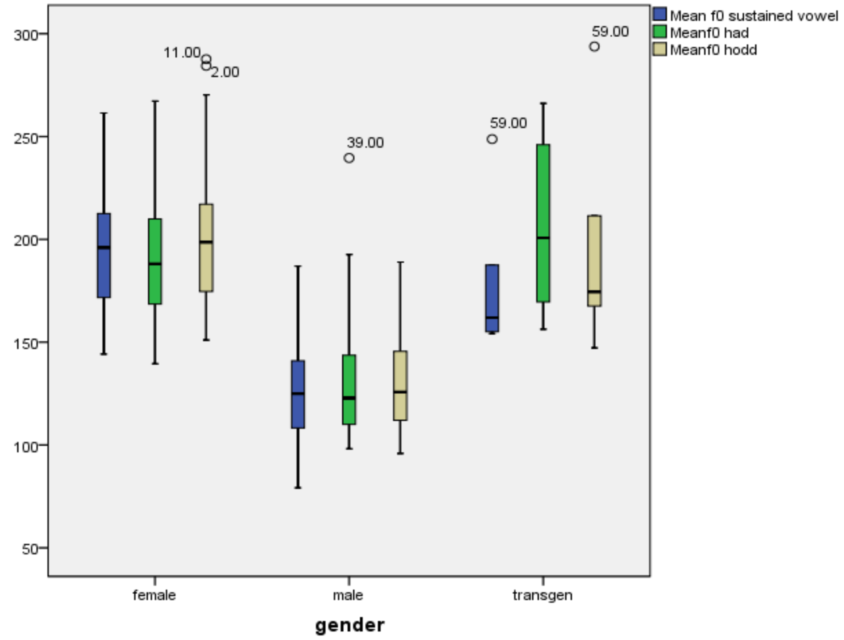
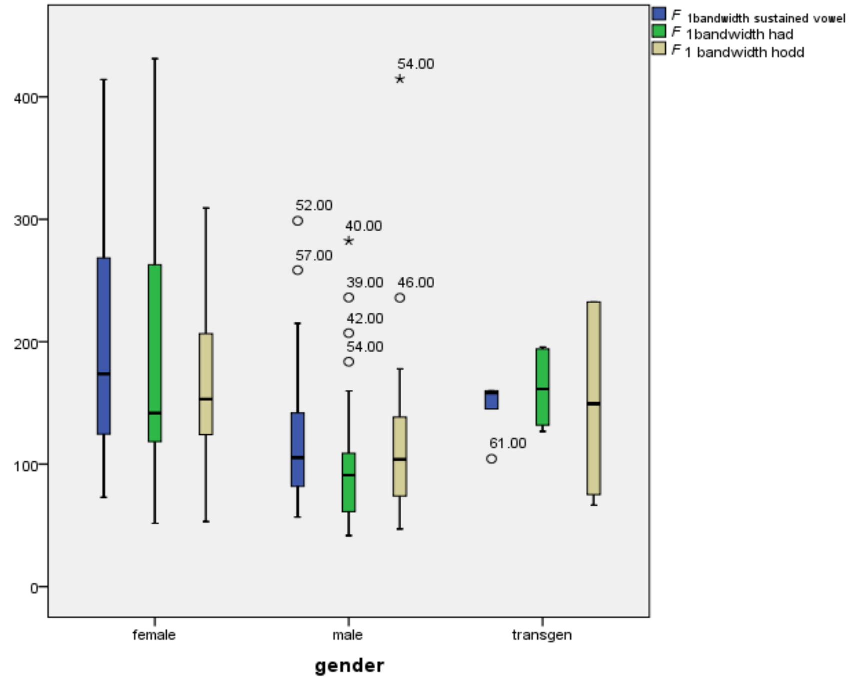


Figure 3: Boxplot of F_1 bandwidth



The third measurement found to be significantly different between men and women across all contexts was CPP: the sustained vowel (*t*=-2.23, *d.f.*=56, *p* <.030; phrase with *had* (*t*=2.97, *d.f.*=56, *p* <.04); and for phrase with *hodd* (*t*=3.301, *d.f.*=56, *p* <.002) As seen in Table 2, the mean for the female voice group was lower than the male voice group. Since a low CPP value has been associated with a high degree of breathiness in the voice, this result may indicate that in the situation of a sustained vowel, the female voice had a greater degree of breathiness than the male voice (Shrivastav and Sapienza, 2003). However, as illustrated in tables 5 and 6, in the context of connected speech the mean CPP of the male voice was found to be lower than the female voices.

The measurement of CPPs followed a similar trend as CPP. A significant difference was found between male and female voices for the sustained vowel (*t*=-6.12, *d.f.* =56, *p* <.000) and mean CPPs for female voices was lower than the male voices. Although no significant difference was found between male and female voices for CPPs measured from connected speech, the mean value of the male voice was lower than the female voice.

Since a low value of CPP and CPPs appears to be related to a female voice when measured from a sustained vowel, the transgender/puberphonic voice group members with the lowest CPP and CPPs values should be identified more often as female. Participants T61 and T63 were found to have the lowest value for CPP and CPPs respectively in the sustained vowel context. As indicated in Tables 8 and 9, a significant difference was found between sexes for solely the measurement of CPP and that in the context of connected speech the lower values of CPP were common to male voices.

These results may indicate that CPP measurement as taken from connected speech is identified in the female voice as a higher value than a male voice; therefore, the individual participant from the transgender/puberphonic voice group with the highest and therefore more closely related to a female value of CPP in this context was T62.

Many other measurements were identified as significantly different between male and female voices but only for certain contexts. Following a similar trend to CPPs, the measurement of H1-H2 was indicated as significantly different only in the sustained vowel context (t=2.4, *d.f.*=56, p <.020). The measurement of H1-A1 was found to be significantly different between sexes for the vowel /ɒ/ in both the sustained (t=3.26, *d.f.*=45.872, p <.002) and sampled from the word *hodd* (t=2.96, *d.f.*=56, p <.005). As seen by Hanson and Chuang (1999), these measurements related to glottal configuration were found to be higher in female voices than male voices in the current results.

Measurements taken by the Multi-Dimensional Voice Profile program from the CSL 4500 (KayPENTAX Corp., 2003) were indicated as significantly different between men and women for only select contexts. Shimmer was indicated as significantly different for both isolated vowels taken from the carrier phrase, /æ/ from *had* (t=3.37, *d.f.*=34.56, p <.002), and /ɒ/ from *hodd* (t=-3.23, *d.f.*=33.04, p <.003). Jitter was observed as significant only in the isolated vowel /ɒ/ from *hodd* (t=-2.42, *d.f.*=30.08, p <.022). Greater values of shimmer and jitter were measured in the male voice than the female voice. Soft phonation index (SPI) was also significant in the isolated vowel /ɒ/ from *hodd* (t=2.57, *d.f.*=56, p <.013). Greater values of SPI were measured in the female voice than male voices for all three contexts.

Table 2: Acoustic measurement *t*-tests with mean and standard deviation of /ɒ/ *sustained vowel.* Note. **p*<.05 level. ***p*<.01 level.

	Gender (Mean ± SD)		
	Female	Male	p
Mean Frequency (Hz)	194.29 ± 29.15	128.60 ± 26.29	.000**
H1-H2 (dB)	4.49 ± 4.30	2.07 ± 3.26	.020*
H1A1 (dB)	5.40 ± 12.41	-3.11 ± 6.87	.002*
H1A3 (dB)	21.21 ± 9.32	17.60 ± 8.09	.122
F_1 bandwidth (Hz)	211.82 ± 117.05	120.96 ± 60.45	.001**
Jitter%	1.06 ± 1.12	0.71 ± 0.33	.105
Shimmer%	3.57 ± 1.89	3.98 ± 1.82	.398
NHR	0.13 ± 0.33	0.14 ± 0.02	.379
VTI	0.04 ± 0.01	0.04 ± .02	.597
SPI	20.73 ± 10.82	17.00 ± 8.85	.158
CPP (dB)	18.80 ± 3.01	20.52 ± 2.89	.030*
CPPs (dB)	7.57 ± 1.30	9.91 ± 1.62	.000**

Table 3**:** Acoustic measurement *t*-tests with mean and standard deviation of /æ/ sampled from *had.* Note. **p*<.05 level. ***p*<.01 level.

	Gender (Mean ± SD)		
	Female	Male	p
Mean Frequency (Hz)	190.83 ± 30.52	132.34 ± 33.26	.000**
H1-H2 (dB)	4.04 ± 3.42	2.70 ± 5.17	.247
H1A1 (dB)	-2.93 ± 5.48	-4.83 ± 6.31	.225
H1A3 (dB)	15.48 ± 6.75	12.07 ± 8.36	.092
F_1 bandwidth (Hz)	222.87 ± 222.81	103.62 ± 59.45	.008**
Jitter%	1.78 ± 1.32	2.43 ± 1.63	.098
Shimmer%	4.87 ± 1.49	7.47 ± 3.82	.002**
NHR	0.30 ± 0.16	0.39 ± 0.19	.070
VTI	0.11 ± 0.06	0.14 ± 0.08	.098
SPI	5.17 ± 2.11	4.65 ± 2.69	.414

Table 4: Acoustic measurement *t*-tests with mean and standard deviation of /ɒ/ sampled from *hodd.* Note. **p*<.05 level. ***p*<.01 level.

	Gender		
	Female	Male	p
Mean Frequency (Hz)	200.46 ± 36.09	131.61 ± 25.83	.000**
H1-H2 (dB)	3.42 ± 2.63	2.04 ± 4.47	.154
H1A1 (dB)	-4.44 ± 4.63	-8.21 ± 5.09	.005**
H1A3 (dB)	17.89 ± 7.98	14.71 ± 9.12	.162
F_1 bandwidth (Hz)	164.74 ± 61.67	119.31 ± 71.79	.012*
Jitter%	1.27 ± 0.80	2.79 ± 3.23	.022*
Shimmer%	4.75 ± 1.76	8.01 ± 5.05	.003**
NHR	0.30 ± 0.20	0.35 ± 0.20	.325
VTI	0.05 ± 0.02	0.07 ± 0.05	.087
SPI	22.33 ± 8.56	16.92 ± 7.30	.013*

For the measurements taken to examine intonation (i.e. semitone range and standard deviation (SD), mean f_0 of the entire phrase and SD, as well as the highest and lowest frequency within the phrase) only those relating to the f_0 and the range of frequencies were significantly different between groups. The f_0 of the entire phrase containing *had* (t=10.68, $d.f.$=56, p <.000) and *hodd* (t=11.16, $d.f.$=56, p <.000), and the SD of f_0 for *had* (t=4.01, $d.f.$=56, p <.000) and for *hodd* (t=3.44, $d.f.$=53.30, p <.001) differed between the sexes, along with the highest (t=5.93, $d.f.$=56, p <.000) and lowest (t=6.40, $d.f.$=56, p <.000) frequencies for the phrase with *had*, and the highest (t=6.48, $d.f.$=56, p <.000) and lowest (t=4.61, $d.f.$=45.28, p <.000) frequencies for the phrase with *hodd*.

Table 5: Acoustic measurement t-tests with mean and standard deviation of the phrase *Please say the word had again.* Note. *p<.05 level. **p<.01 level.

	Gender (mean ± SD)		
	Female	Male	p
Mean f_0 phrase (Hz)	190.39 ± 21.98	128.86 ± 21.92	.000**
SD f_0 phrase (Hz)	38.32 ± 15.42	23.75 ± 11.87	.000**
Highest frequency (Hz)	288.07 ± 50.55	199.95 ± 61.95	.000**
Lowest frequency (Hz)	128.01 ± 22.48	94.52 ± 16.69	.000**
Semitone range	14.03 ± 4.59	12.50 ± 5.40	.248
SD of semitones	3.27 ± 1.09	2.93 ± 1.23	.265
CPP (dB)	12.80 ± 0.98	12.07 ± 0.90	.004*
CPPs (dB)	4.38 ± 0.51	4.21 ± 0.52	.224

Table 6: Acoustic measurement t-tests with mean and standard deviation of phrase *Please say the word hodd again.* Note. *p<.05 level. **p<.01 level.

	Gender (mean ± SD)		
	Female	Male	p
Mean f_0 (Hz)	191.55 ± 22.05	128.59 ± 20.86	.000**
SD f_0 phrase (Hz)	36.20 ± 17.53	22.35 ± 12.97	.001**
Highest frequency (Hz)	275.68 ± 47.91	188.20 ± 54.87	.000**

	Gender (mean ± SD)		
	Female	Male	p
Lowest frequency (Hz)	124.47 ± 30.72	94.84 ± 16.63	.000**
Semitone range	14.13 ± 5.24	11.46 ± 5.30	.059
SD of semitones	3.15 ± 1.33	2.77 ± 1.34	.280
CPP (dB)	12.80 ± 0.97	12.01 ± 0.87	.002*
CPPs (dB)	4.43 ± 0.55	4.22 ± 0.55	.172

The acoustic measurements, which were found to be significantly different between men and women, taken from the transgender group participants were standardized as *z*-scores by comparison to the female group mean and standard deviation. These results are presented in tables 7, 8, and 9. The majority of the transgender group participants fell within 2 standard deviations of the female mean; however, the mean f_0 taken from the recording of carrier phrases produced by T59 was found to be above the female mean by 2 standard deviations. Participant T60 had a F_1 bandwidth value well above the female group mean during production of the carrier phrase with *hodd*. Participants T60 and T62 were found to produce CPPs values 3 standard deviations above the female mean when producing the sustained vowel. Although above the mean of the female group, these participants are well beyond the expected male values for these measurements, thus these individuals may be more successful in being perceived as a female than other members of the group.

T61 voice production when reading the carrier phrase with *had*, resulted in a shimmer percentage just above 2 standard deviations and a mean f_0 taken from the entire phrase 2 standard deviations below the female group mean. Converse to the data taken from T59, T60, and T62, due to the relationship seen in the data compared between male and female voices – e.g. greater percentage of shimmer and lower f_0 in the male voice than female voice - these results may show that T61 will be less successful in achieving feminine perception than other participants in the group.

Overall several acoustic measurements appeared to be significantly different between men and women depending on the speaking context recorded. The measurements that were found to differ significantly between male and female voices across contexts were f_0, F_1 bandwidth and CPP. Successful portrayal of the female voice and therefore perception as a female for the members of the transgender group maybe dependent on several of the acoustic measurements found to be significantly different between men and women. The results for listener perception of gender are presented in the next section.

Table 7: z-scores of acoustic measurements, found to be significantly different between men and women taken from the sustained vowel produced by transgender/puberphonic group participants

Context	*z*-scores				
	T59	T60	T61	T62	T63
Mean Frequency (Hz)	1.87	-0.23	-1.34	-1.37	-1.11
H1-H2 (dB)	-0.19	0.29	-0.30	0.17	-0.72
H1A1 (dB)	-0.22	-0.51	-0.61	-0.62	-0.42
F_1 bandwidth (Hz)	-0.57	-0.44	-0.92	-0.46	3.72
CPP (dB)	0.13	0.73	-0.75	1.41	-1.28
CPPs (dB)	-1.14	3.70	-1.36	3.05	0.16

Table 8: z-scores of acoustic measurements, found to be significantly different between men and women for carrier phrase: /æ/ sampled from *had* + intonation measurements produced by transgender/puberphonic group participants

Context	*z*-scores				
	T59	T60	T61	T62	T63
Mean Frequency (Hz)	2.467871	1.808933	-0.69849	0.324585	-1.13158
F_1 bandwidth (Hz)	-0.12204	-0.40829	-0.27597	-0.12911	-0.43122
Shimmer%	-0.09787	-0.85569	2.711914	0.674052	1.082833
CPP (dB)	-0.82682	-0.55156	-1.42833	-0.29668	-1.06131
Mean f_0 phrase (Hz)	2.239361	1.329132	-2.10145	-0.35354	-1.64383
SD f_0 phrase (Hz)	-0.38456	-0.04993	-1.04018	0.110244	-1.82356
Lowest frequency (Hz)	1.819966	-0.68486	-0.66217	-1.06029	0.518402
Highest frequency (Hz)	1.076151	-0.26398	-1.89452	-0.65232	-1.71054

Table 9: z-scores of acoustic measurements, found to be significantly different between men and women for carrier phrase: /ɒ/ sampled from *hodd* + intonation measurements produced by transgender/puberphonic group participants

Context	*z*-scores				
	T59	T60	T61	T62	T63
Mean Frequency (Hz)	2.585975	0.303908	-0.72006	-0.91077	-1.47428
H1A1 (dB)	1.084996	1.871824	0.776333	-0.11605	0.43648
F_1 bandwidth (Hz)	1.097128	9.924274	-1.59441	-0.2497	-1.45306
Jitter%	-0.32075	-0.14188	-0.0956	-0.43207	-0.44457
Shimmer%	-1.28092	-1.24673	0.955396	0.58619	0.869932
SPI	0.156544	-1.15079	0.346081	0.467244	-0.633
CPP (dB)	-1.55896	-1.03381	-0.65283	-0.18946	-1.30153
Mean f_0 phrase (Hz)	1.870536	0.775975	-1.72659	-1.527	-1.89942
SD f_0 phrase	0.226756	-0.2816	-0.83333	0.121775	-1.64978
Lowest frequency	1.560715	0.254241	-0.36953	-0.36953	-0.00295
Highest frequency	0.768545	-0.37469	-1.33216	-0.42959	-1.97297

4.3 LISTENER MEASUREMENTS

Despite the fact that several of the acoustic measurements, which were found to differ significantly between men and women, were exhibited within 2 standard deviations of the female mean in the participants of the transgender group, listeners often agreed with a high majority that voice samples produced by the transgender group were produced by men. Listeners also rated members of the transgender group as mildly masculine, especially in connected speech production.

For each context listeners were required to complete two tasks for each voice heard: first, to choose which gender, either male or female, he or she believed the speaker was; and secondly, to rate the masculinity or femininity, as it corresponded to the gender chosen, of that voice. To determine which gender – i.e. male or female - the majority of the listeners judged each voice as, a percent correct identification of gender or percent identification of gender was calculated.

Percent 'correct' identification of gender was based on the expected listener perception of the voice groups – i.e. female perception for the biological female and transgender groups, and male perception of the male group. The number of listeners judging a voice participant as male and the number of listeners judging a voice participant as female were tallied. The number of 'correct' gender judgements – e.g. female for biological female or transgender group members – was divided into the number of total judgements. The complete list of response data is provided in appendix 3 for each listening context. The percent identification of gender remained consistent across contexts for the male voice recordings (sustained vowel gender identification = 93.41%, phrase *had* gender identification = 96.98%, phrase *hodd* gender identification = 97.25%). For female recorded voices the percent gender identification remained fairly consistent between the two carrier phrases (phrase *had* gender identification = 98.97%, phrase *hodd* gender identification = 98.72%); however, decreased for the percent gender identification of the sustained vowel context (65.81%). The transgender group was similar to the female group in that the participants were identified inconsistently among contexts (sustained vowel gender identification = 48.46%, phrase *had* gender identification = 28.46%, phrase *hodd* gender identification = 17.69%).

The perception of gender appears to be context dependent, especially for the female and transgender group. Consistency of perception of gender by context influenced the reliability of each listener as well. Interrater reliability for listener judgement in the dichotomous choice of gender - i.e. female versus male - was determined using Fleiss's Kappa, as there were multiple raters (N=26) of the nominal data for each voice recorded. Overall Kappa for sustained vowel was, $\kappa = 0.62$ (SE=0.004, 95% CI=0.61 to 0.62), for the carrier phrase with *had* was, $\kappa = 0.91$ (SE=0.007, 95% CI=0.90 to 0.92), and for the carrier phrase with *hodd* was, $\kappa = 0.92$ (SE=0.007, 95% CI=0.90 to 0.93). Despite changing among contexts, interrater reliability for the dichotomous choice of speaker gender ranged from strong to near complete agreement across contexts (Fleiss, 1985 as cited by Chang, 2011).

The interrater reliability for listener scaled rating of masculinity/femininity as indicated from two scales (i.e. masculinity 0 to 100; femininity 0 to -100) was high across vowels and contexts: the inter rater reliability yielded an Intraclass Correlation Coefficient (ICC) of 0.99 for the phrase with *had*, 0.97 for the phrase with *hodd* and 0.94 for the sustained vowel.

Intrarater reliability was calculated for each listening participant using the sustained vowel production. Unfortunately, carrier phrases were played only once for listeners and therefore intrarater reliability could not be calculated for these contexts. For the sustained vowel production, the listener heard three segments from the same recording of a sustained vowel sample for each participant during the perception experiment. For the choice of speaker gender Fleiss kappa was used for the three rated segments for each listener. The overall kappa ranged from 0.41 to 0.89 with a mean of 0.72. This would indicate a strong mean intrarater reliability for gender perception (Fleiss, 2010). The same three segments were used to determine intrarater reliability for the gender score (i.e. masculinity/femininity) for each listener. The ICC yielded an ICC range of 0.72 to 0.96 with a mean ICC of 0.91.

As shown in the results from percent identification and listener reliability, when given greater amounts of acoustic information – i.e. listening to connected speech versus a sustained vowel for approximately 400ms – distribution of gender perception changed. In other words, men and women were more often identified as their actual sex and transgender/puberphonic voice group members were identified as their biological sex when more acoustic information was provided – i.e. listening to the carrier phrases. However, when the amount of time and information provided to the listener was reduced, as in the context of the 400ms portions of the sustained vowel, the certainty of perceiving the female speaker and the transgender/puberphonic voice group speaker decreased – i.e. the perception of gender became increasingly evenly distributed between male or female. Conversely, perception of the male speakers remained fairly accurately distributed in this situation, indicating a possible overall bias towards reporting a voice as male when less acoustic information is provided.

As noted previously, many of the individual participants in the transgender group were not overly successful in the perception of their voice as female or feminine by listeners. Percentage of listeners who identified individuals from the transgender/puberphonic group recordings as female was consistently low across contexts. Looking at the individual results of the group, as found in table 10, participant T59 was consistently rated higher on the scale of femininity than the other participants. This may implicate the importance of f_0 to gender perception because f_0 is indicated as significantly different between men and women, and that T59 produced the highest value for the transgender/puberphonic voice group of mean f_0 consistently. For the results of the sustained vowel context shown in table 10, the transgendered speaker with the next highest percentage identification of gender for "female" was participant T63. Interestingly, T63 did not produce the highest or even second highest f_0 for the transgender/puberphonic voice group. T63, as identified in table 9, did produce the highest F_1 bandwidth and the lowest CPP value, a possible indication of the importance of those measurements in the perception of gender for this context.

Even within the context of connected speech, listeners' perception of gender differed. Results that are presented in Table 11 show that participant T59 and participant

T62 had the highest percent of listeners who identified the participants as female (57.69%) in the phrase with *had*. T59 was rated as more feminine (-21.42/100) than T62 (-16.81/100). Participant T62 had only the third highest f_0 of the transgender/puberphonic voice group, once again possibly indicating an interaction of f_0 with other measurements in the listener perception of gender. T62 was identified as following the trend of the female voice for producing a greater value of CPP in connected speech and a greater standard deviation of frequency than the male voice.

Surprisingly, in the carrier phrase with *hodd*, participant T63 was identified as female with the highest percent gender identification of listeners (38.46%), followed by participant T59 (34.62%). However, participant T59 was still rated highest for perception of femininity, followed by T63, meaning that although listeners agreed less often that T59's voice sample was produced by a female, the participant still was given a higher feminine rating on average than T63. As seen in table 12, participant T63 had an average gender rating of 1.54 out of 100 on the masculine scale for this context. Listeners may most often perceive T63's voice sample as mildly masculine although judging her voice as produced by a female with 38.46% gender identification among listeners. Participant T63 in the context of the carrier phrase with *hodd,* had the lowest f_0 of the transgender/puberphonic voice group and within the male range. Nevertheless, T63 was indicated with the highest percent gender identification of listeners as a female. This result may be attributed to an interaction between different acoustic measurements or to vocal/speech elements yet to be explored.

Table 10: Listeners response and average feminine/masculine score for sustained vowel segments

Participant	segment	Female	Male	Percent female	Average rating
T59	1	26	0	100.00	-72.73
T59	2	23	3	88.46	-64.73
T59	3	24	2	92.31	-54.73
T60	1	9	17	34.62	7.65
T60	2	3	23	11.54	18.96
T60	3	5	21	19.23	14.42
T61	1	6	20	23.08	24.88
T61	2	3	23	11.54	23.88
T61	3	1	25	3.85	25.08
T62	1	3	23	11.54	41.31
T62	2	6	20	23.08	20.85
T62	3	15	11	57.69	-1.81
T63	1	22	4	84.62	-22.77
T63	2	22	4	84.62	-25.77
T63	3	21	5	80.77	-25.88
Mean				48.46	-6.09
SD				±36.19	±36.21

Table 11: Listeners response and average feminine/masculine score for carrier phrase with *had*

Participants	Female	Male	Percent female	Average rating
T59	15	11	57.69	-21.42
T60	0	26	0.00	11.12
T61	0	26	0.00	50.38
T62	15	11	57.69	-16.81
T63	7	19	26.92	10.04
Mean			28.46	6.66
SD			±28.86	±28.67

Table 12: Listeners response and average feminine/masculine score for carrier phrase with *hodd*

Participants	Female	Male	Percent female	Average rating
T59	9	17	34.62	-10.35
T60	2	24	7.69	11.15
T61	0	26	0.00	36.77
T62	2	24	7.69	31.23
T63	10	16	38.46	1.54
Mean			17.69	14.07
SD			±17.54	±19.82

In exploring less successful portrayal of the expected gender by participants, listeners perceived participant T61 as male with the highest percent gender identification among listeners across all three contexts. Information from the previous section in this chapter predicted this result for T61. Despite participant T61 having a low $f_{0,}$ she had a mean f_0 within the gender ambiguous range. This fact combined with the result of listener perception appears to be an indication that other cues important to the listener in identifying a female voice were lacking. The transgender participants with low percentage gender identification of perception of female gender were T60, T61 and T62 for the sustained vowel. When compared with the male speakers both T61 and T62 were rated within one standard deviation of the mean of masculinity rating. Participant T60 had the lowest masculinity rating (13.68/100) of those transgender participants more often perceived as male.

Five of the female participants (F8, F10, F21, F24, and F29) were perceived with the highest listener percent gender identification as male voices. These voices were individually reported as one standard deviation below the average percent gender identification of listeners identifying a voice as female for the sustained vowel. One female participant, F24, was indicated as male with a percent gender identification of listeners two standard deviations below the female mean consistently across the three segments of the sustained vowel. She was infrequently identified as a female. Her voice was perceived by only one listener as female in two out of the three portions of the sustained vowel, and perceived by all listeners as male in the final segment (0-3.85%). In contrast, listeners reported with 100% gender identification F24 as a female for both carrier phrases; however, femininity rating of the voice was perceived as only moderately high (-57.27 and -64.08) in this context. She was rated as moderately masculine (25-50) at 34.17 for the production of the sustained vowel: one standard deviation away from the mean for female recorded voices indicating increasing masculinity. As provided in table 13, the individual data of F24 indicates a fairly consistent mean f_0, although slightly lower than the female group mean. Since F24 was perceived more often as female than male from the connected speech sample, this may indicate that greater amounts of information and possibly different cues to the listener are available in connected speech versus sustained vowel.

Table 13: Individual data with *z*-scores based on female group means for participant F24

F24 measurements (*z*-score)				
	/ɒ/	/æ/ had	/ɒ/ hodd	Unit
f_0	172.89 (-0.73)	174.03 (-0.55)	178.99 (-0.59)	Hz
H1-H2	6.42 (0.45)	1.06 (-0.87)	1.72 (-0.65)	dB
H1-A1	-0.25 (-0.46)	-3.23 (-0.05)	-6.66 (-0.48)	dB
H1-A3	26.30 (0.55)	8.81 (-0.99)	13.48 (-0.55)	dB
$F1$ BW	114.03 (-0.84)	84.75 (-0.62)	124.13 (-0.66)	Hz
% jitter	0.40 (-0.59)	1.1 (-0.53)	1.1 (-0.22)	%
% shimmer	2.88 (-0.37)	3.33 (-1.04)	5.19 (0.25)	%
NHR	0.12 (-0.51)	0.22 (-0.53)	0.18 (-0.58)	
VTI	0.033 (-0.68)	0.14 (0.57)	0.078 (1.46)	
SPI	41.30 (1.90)	4.41 (-0.36)	18.74 (-0.42)	
CPP	19.38 (0.19)			dB
CPPs	8.89 (1.02)			dB
Phrase:		Had	Hodd	
CPP		19.38 (1.02)	12.75 (-0.06)	dB
CPPs		5.01 (1.25)	4.42 (-0.01)	dB
Mean f_0		152.90 (-0.84)	167.12 (-0.80)	Hz
f_0 SD		38.21 (-1.49)	26.25 (-1.13)	Hz
Semitones		6 (-1.75)	8 (-1.17)	
SD semitone		1.5 (-1.63)	1.56 (-1.19)	
% Gender identification				
Female	0.67	26	26	Counts
Male	25.33	0	0	Counts
Proportion	2.56	100	100	%
Average rating	34.17	-57.27	-64.08	

Female participants had an overall high percent gender identification of listeners identifying the voices as female when heard speaking the carrier phrases; however, within this group individual average femininity scores varied. Across the two phrases, participants F9 and F29 were rated by listeners as one standard deviation below the mean for the female group. For the phrase containing *had*, F9 achieved percent gender identification of listeners of 96.1% but with a low femininity rating of -25.3 out of -100. For the carrier phrase with *hodd*, this participant achieved a high percent gender identification of listeners perceiving the voice as female but a low femininity score, relative to the other participants of the female voice group, of -30.3 out of -100. For participant F29, her percent gender identification remained at 84.6% listener agreement as the voice as female across the two phrases and her scores for the phrases containing *had* and then *hodd* were -27.7 and -25.1 out of -100.

Only one male voice, participant M33, was recorded as having an especially low percent gender identification for perception as male, which remained consistent between the two carrier phrase contexts (phrase *had* gender identification = 34.62% and phrase *hodd* = 30.77%). The percentage gender identification for this participant increased to 91.03% for the sustained vowel production. Participant M44 (represented as T59 in the transgender group in his puberphonic voice), was the only male participant consistently indicated by listeners as male with a percent gender identification one standard deviation below the mean for the sustained vowel production of the male group (53.85-73.08%). M44 data indicated 100% listener gender identification for male gender in the carrier phrase with *hodd*, and 88.46% for carrier phrase with *had* both perceived as a male. His individual data for both his puberphonic voice (T59) and male voice (M44) is listed in table 14.

Table 14: Individual data with *z*-scores based on male group mean of the male voice of participant M44, and the individual data of the puberphonic voice of participant T59

Acoustic measurements							
	M44 (*z*-scores)			T59			
	/ɒ/	/æ/ had	/ɒ/ hodd	/ɒ/	/æ/ had	/ɒ/ hodd	Unit
f_0	168.36 (1.51)	164.17 (0.96)	134.18 (0.10)	248.7	266.14	293.77	Hz
H1-H2	2.66 (0.18)	4.64 (0.38)	5.22 (0.71)	3.68	7.40	7.93	dB
H1-A1	13.59 (2.43)	-6.60 (-0.28)	-8.10 (0.02)	2.72	1.64	0.58	dB
H1-A3	31.03 (1.66)	15.93 (0.46)	9.11 (-0.62)	28.57	27.49	38.83	dB
*F*1 BW	104.89 (-0.27)	96.64 (-0.12)	112.8 (-0.09)	145.1	195.68	232.40	Hz
% jitter	0.94 (0.69)	1.42 (-0.62)	1.58 (-0.38)	0.42	1.38	1.02	%
% shimmer	4.06 (0.04)	6.56 (-0.24)	4.17 (-0.76)	2.94	4.73	2.51	%
NHR	0.14 (-0.16)	0.67 (1.48)	0.31 (-0.18)	0.11	0.46	0.15	
VTI	0.05 (0.51)	0.21 (0.87)	0.11 (0.77)	0.04	0.06	0.04	
SPI	18.63 (0.18)	4.02 (-0.23)	11.39 (-0.76)	16.71	3.78	23.67	
CPP	16.63 (-1.35)			19.19			dB
CPPs	9.68 (-0.14)			6.08			dB
Phrase:		Had	Hodd		Had	Hodd	
CPP		11.44	11.51 (-0.57)		11.99	11.29	dB
CPPs		3.59	3.52 (-1.27)		2.49	4.04	dB
Mean f_0		145.61	135.27 (0.32)		239.47	232.79	Hz
f_0 SD		29.15	20.14 (-0.17)		32.39	40.17	Hz
Semitones		21 (1.57)	12 (0.10)		13.00	10.00	
SD semitone		3.33 (0.32)	2.51 (-0.19)		2.31	2.78	
% Gender identification							
Female	10.33	3	0	24.33	15	9	Counts
Male	15.67	23	26	1.67	11	17	Counts
Proportion	60.26	88.46	100	93.59	57.69	34.62	%
Average rating	2.35	17.27	35.19	-64.1	-21.42	-10.35	dB

Participant M33 and M44 were below one standard deviation of the mean for masculinity rating of the male group (*had* mean=65.41, *hodd* mean=65.98, sustained vowel=64.48) indicating a low masculinity rating, consistently over all three conditions. Although misidentified as female, participant M33 was rated as only mild (0 to -25) to moderately feminine (-26 to -50) for the phrase conditions. As well, while correctly identified as male, M44 was rated as mild to moderately masculine. Individual data for M33 is listed in table 15.

Table 15: Individual data with *z*-scores based on male group mean of participant M33

	M33 acoustic measurements (*z*-score)			Unit
	/ɒ/	/æ/ had	/ɒ/ hodd	
f_0	164.36 (1.36)	192.58 (1.81)	188.96 (2.22)	Hz
H1-H2	3.22 (0.35)	4.18 (0.29)	2.40 (0.08)	dB
H1-A1	-7.11 (-0.58)	-12.67 (-1.24)	-13.75 (-1.09)	dB
H1-A3	10.48 (-0.88)	-3.92 (-1.91)	2.76 (-1.31)	dB
F_1 BW	80.32 (-0.67)	79.49 (-0.41)	94.93 (-0.34)	Hz
% jitter	0.97 (0.80)	1.20 (-0.75)	0.98 (-0.56)	%
% shimmer	2.60 (-0.76)	3.06 (-1.16)	0.46 (-1.49)	%
NHR	0.13 (-0.58)	0.52 (0.72)	0.37 (0.09)	
VTI	0.03 (-1.06)	0.07 (-0.84)	0.06 (-0.14)	
SPI	10.90 (-0.69)	2.48 (-0.80)	11.02 (-0.81)	
CPP	22.32 (0.62)			dB
CPPs	8.74 (-0.14)			dB
Phrase:		Had	Hodd	
CPP		12.23 (0.18)	12.46 (0.52)	dB
CPPs		4.36 (0.29)	4.44 (0.39)	dB
Mean f_0		170.99 (1.92)	174.74 (2.21)	Hz
f_0 SD		24.57 (0.07)	27.13 (0.37)	Hz
Semitones		6 (-1.20)	8 (-0.65)	
SD semitone		2.30 (-0.51)	2.65 (-0.09)	
% Gender identification				
Female	2.33	17	18	Counts
Male	23.67	9	8	Counts
Proportion	91.03	34.62	30.77	%
Average rating	25.14	-24.27	-29.12	

The transgender participants as well as the one individual with puberphonia varied greatly in listener percent gender identification. One transgender participant (T60) provided her male voice (represented as M46) as well as her female voice. Listeners identified her voice as male reliably even when hearing her female vocal production within the female range of f_0. Participant T60's male voice received a gender score range from moderate (45.12) to moderately high (72.12) in masculinity. Her female voice was perceived as masculine as well, although her gender score was only mildly masculine (11.12 to 13.68). The comparison of her data as male to transgender female is provided in table 16.

Table 16: Individual data with *z*-scores based on male group mean of the male voice of participant M46, and the individual data of the female voice of participant T60

Acoustic measurements							
	M46 (*z* scores)			T60			
	/ɒ/	/æ/ had	/ɒ/ hodd	/ɒ/	/æ/ had	/ɒ/ hodd	Unit
f_0	121.21 (-0.28)	126.4 (-0.18)	135.90 (0.17)	187.5	246.03	211.42	Hz
H1-H2	4.96 (0.89)	0.37 (-0.45)	2.70 (0.15)	5.73	13.19	8.28	dB
H1-A1	-3.81 (-0.10)	-7.08 (-0.36)	-8.58 (-0.07)	-0.97	-2.88	4.22	dB
H1-A3	18.05 (0.06)	3.92 (-0.98)	8.84 (-0.64)	12.08	10.17	18.33	dB
*F*1 BW	129.22 (0.14)	129.43 (0.43)	235.88 (1.62)	160.0	131.91	776.78	Hz
% jitter	0.27 (-1.31)	2.48 (0.03)	1.57 (-0.38)	0.22	2.97	1.16	%
% shimmer	3.29 (-0.38)	15.53 (2.11)	8.63 (0.12)	1.34	3.60	2.57	%
NHR	0.12 (-1.14)	0.72 (1.76)	0.54 (0.96)	0.12	0.19	0.22	
VTI	0.008 (-2.26)	0.33 (2.44)	0.06 (-0.14)	0.04	0.06	0.07	
SPI	16.63 (-0.04)	2.35 (-0.85)	14.17 (-0.38)	40.70	5.17	12.43	
CPP	24.44 (1.36)			21.00			dB
CPPs	12.93(1.87)			12.38			dB
Phrase:		Had	Hodd		Had	Hodd	
CPP		12.69 (0.69)	12.28 (0.32)		12.26	11.80	dB
CPPs		4.52 (0.60)	4.32 (0.17)		3.92	3.32	dB
Mean f_0		132.67 (0.17)	135.21 (0.32)		219.52	208.66	Hz
f_0 SD		19.32 (-0.37)	19.30 (-0.24)		37.55	31.26	Hz
Semitones		10 (-0.46)	9 (-0.46)		16.00	12.00	
SD semitone		2.49 (-0.36)	2.48 (-0.22)		3.39	2.83	
% Gender identification							
Female	0	0	0	5.67	0	2	Counts
Male	26	26	26	20.33	26	24	Counts
Proportion	100	100	100	20.33	0	7.69	%
Average rating	72.12	49.12	45.12	13.68	11.12	11.15	dB

The results of the listener response identifying gender indicate that across contexts f_0 more than likely plays an important role in gender perception. Nevertheless, interaction between other elements of the voice and speech may also have a meaningful role as evident from the perception of some voices within a female f_0 range as masculine. By formulating a model for prediction of gender perception using the listener response data and the acoustic measurements will inform the development of gender perception and the voice. Results may inform speech-language pathologists about the most important acoustic properties that cue gender perception in listeners, so that they can focus on manipulating those acoustic properties in voice therapy with MtF transgender clients.

4.4 VOICE QUALITY CORRELATIONS OF ACOUSTIC MEASUREMENTS AND GENDER PERCEPTION

A multiple stepwise regression was performed for each context to examine the relationship between components of the acoustic measurements analyzed from the voice samples from all participants and the listener ratings of masculinity/femininity. For all three contexts mean fundamental frequency (f_0) was indicated as a significant predictor of perception of gender (all three contexts p <.000). The data of mean f_0 for the three different contexts and perception of gender is illustrated in the scatterplots in figure 4, 5 and 6. For each context the combination of measurements that account for the greatest variance in perception of gender differed.

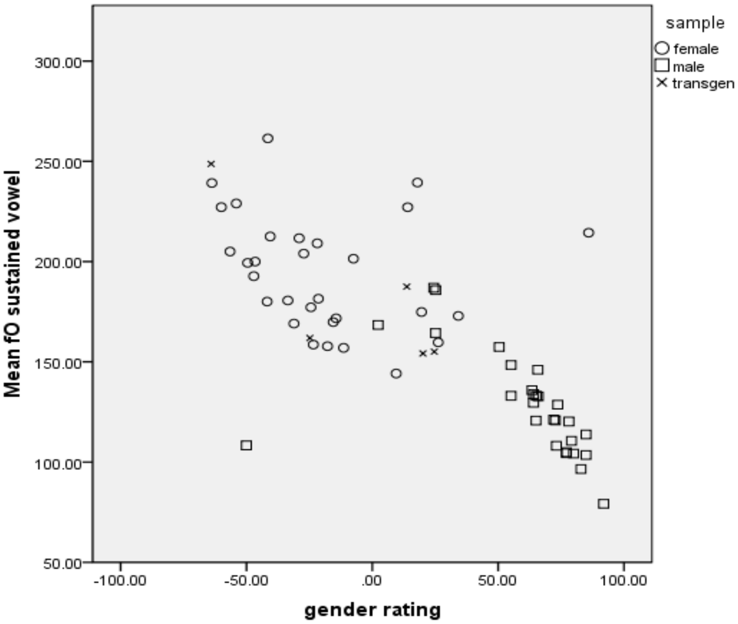


Figure 4: Scatterplot of male, female and transgender participants as perceived by listener response to masculinity (100)/femininity (-100) scores compared to fundamental frequency from the sustained vowel production

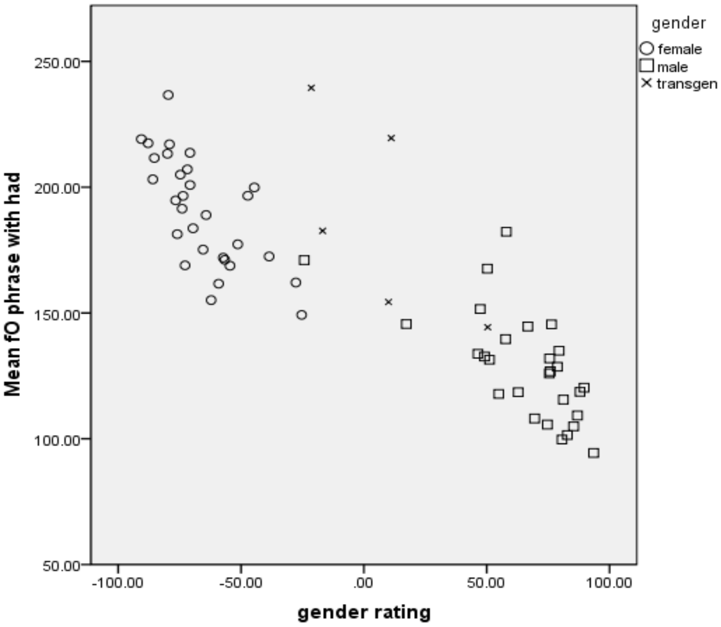


Figure 5: Scatterplot of male, female and transgender participants as perceived by listener response to masculinity (100)/femininity (-100) scores compared to fundamental frequency from the carrier phrase: *Please say the word had again.*

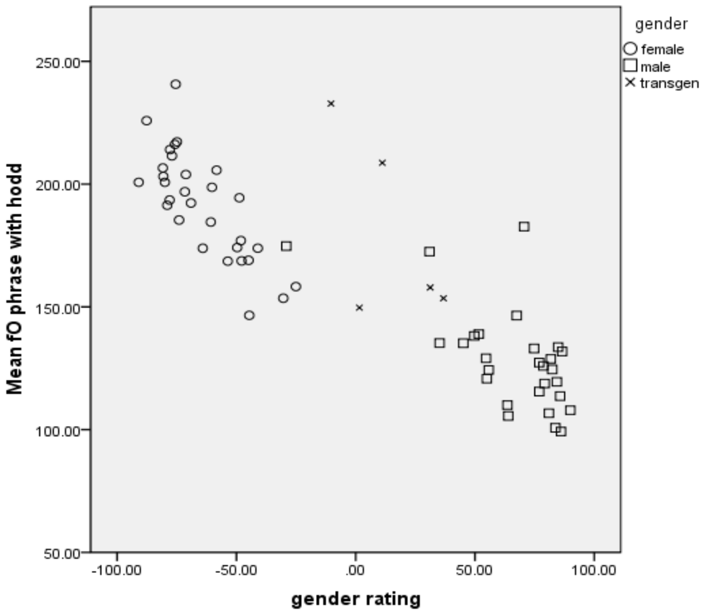


Figure 6: Scatterplot of male, female and transgender participants as perceived by listener response to masculinity (100)/femininity (-100) scores compared to fundamental frequency from the carrier phrase: *Please say the word hodd again.*

In the sustained vowel context when combined with f_0, the acoustic measurements accounting for the greatest variance in perceptual ratings were voice turbulence index (VTI), H1-A3 and soft phonation index (SPI). This combination of acoustic properties accounted for 67.0% of variance in the listeners' ratings of masculinity/femininity for the sustained vowel.

Table 17: Stepwise regression predicting masculinity/femininity ratings from acoustic measures for sustained vowel /ɒ/

Sustained vowel	R	Adjusted R-squared	F (degrees of freedom)	Significance
Mean f_0	.753	.561	80.10 (1, 61)	.000
f_0 + VTI	.792	.615	50.56 (2, 60)	.000
f_0 + VTI + H1-A3	.817	.650	39.42 (3, 59)	.000
f_0 + VTI + H1-A3 + SPI	.832	.670	32.53 (4, 58)	.000

For the acoustic measurements taken from the isolated vowels from *had* and *hodd*, the combination of f_0 with H1-H2 measurements accounted for 57.9% and 63.2% respectively, of variance in listener's perceptual ratings. The combination of f_0, H1-H2 and percentage of jitter accounted for 64.5% of the variance of listener perception when listening to the carrier phrase with *hodd*.

Table 18: Stepwise regression predicting masculinity/femininity ratings from acoustic measures for isolated vowel from *had* /æ/

Had	R	Adjusted R-squared	F (degrees of freedom)	Significance
Mean f_0	.709	.494	61.57 (1, 61)	.000
f_0 + H1-H2	.761	.565	41.34 (2, 60)	.000

Table 19: Stepwise regression predicting masculinity/femininity ratings from acoustic measures for isolated vowel from *hodd* /ɒ/

Hodd	R	Adjusted R-squared	F (degrees of freedom)	Significance
Mean f_0	.759	.569	82.75 (1, 61)	.000
f_0 + H1-H2	.795	.619	51.43 (2, 60)	.000
f_0 + H1-H2 + jitter	.814	.645	38.55 (3, 59)	.000

In contrast, with the addition of the variables taken from the whole phrase - i.e. mean f_0 of carrier phrase, CPP and CPPs - the combination of mean f_0 and CPP taken from connected speech accounted for 79.9% of variance in the phrase with *hodd* and 76.7% in phrase with *had*.

Table 20: Stepwise regression predicting masculinity/femininity ratings for acoustic measurements for the carrier phrase: *Please say the word had again*

Had	R	Adjusted R-squared	F (degrees of freedom)	Significance
Mean f_0 entire phrase	.852	.726	161.7 (1,61)	.000
f_0 + CPP	.876	.767	98.8 (2, 60)	.000

Table 21: Stepwise regression predicting masculinity/femininity ratings for acoustic measurements for the carrier phrase: *Please say the word hodd again*

Hodd	R	Adjusted R-squared	F (degrees of freedom)	Significance
Mean f_0 entire phrase (Hz)	.868	.749	186.2 (1, 61)	.000
f_0 + CPP	.898	.799	124.5 (2, 60)	.000

When all measurements of the two carrier phrases were combined the combination of mean f_0 with H1-H2 and CPPs accounted for the highest amount of variance (79.2%) for the phrase with *had*, and a combination of f_0, CPP and H1-H2 accounted for 81.9% of variance for the phrase with *hodd*.

Table 22: Stepwise regression predicting masculinity/femininity ratings for all acoustic measurements for the carrier phrase: *Please say the word had again*

Had	R	Adjusted R-squared	F (degrees of freedom)	Significance
Mean f_0 entire phrase (Hz)	.852	.722	161.7 (1, 61)	.000
f_0 + H1H2	.884	.774	106.9 (2, 60)	.000
f_0 + H1H2 + CPPs	.895	.792	79.5 (3, 59)	.000

Table 23: Stepwise regression predicting masculinity/femininity ratings for all acoustic measurements for the carrier phrase: *Please say the word hodd again*

Hodd	R	Adjusted R-squared	F (degrees of freedom)	Significance
Mean f_0 entire phrase (Hz)	.868	.749	186.2 (1, 61)	.000
f_0 + CPP	.898	.799	124.5 (2, 60)	.000
f_0 + CPP + H1H2	.910	.819	94.8 (3, 59)	.000

Largely the measurements found to be part of a significant predictive model of gender perception when combined with the primary indicator, f_0, differed between the sustained vowel context and the carrier phrases. It appears that a predictive model for connected speech includes H1-H2 amplitude relations, and the cepstral peak prominence measurements. Further variation exists within the isolated vowels from the carrier phrases with the addition of percentage of jitter to the predictive model in the carrier phrase with *hodd*; however, the highest variance was accounted for when the acoustic components measuring the entire phrase were added to the model.

A different group of acoustic properties emerges when the listener hears short portions of a sustained vowel. In this context, the addition of two of the measurements associated with noise parameters in the signal, SPI and VTI, and H1-A3 amplitude relations create a significant predictive model of gender perception when combined with f_0.

CHAPTER 5: DISCUSSION

5.1 GOALS OF THE STUDY

Training the voice to match the gender identity of the individual is an essential part of achieving acceptance in a desired gender, which is necessary to the well-being of a transgender individual. Whether with or without assistance from a speech-language pathologist, a male-to- female transgender individual may alter the voice as she alters her physical appearance, by way of different acoustic parameters she believes are feminine - i.e. intonation or fundamental frequency; however, the perception of femininity is likely more complicated than just pitch and intonation alone. The different acoustic parameters society deems appropriate and identifiable to each gender are individually complex as well as multifarious in the interaction between them. The current study investigated acoustic measurements related to pitch, intonation, and voice quality, and the relationship between these measurements and perception of gender. Several measures were shown to differ significantly between gender groups and to hold important relationships with listeners' perception of gender differences.

The first question posed by this study was: what acoustic measurements help discriminate between male and female voices? Voice parameters known to be relevant to gender perception were compared between sexes and among the three gender groups (male, female and transgender). Three measurements were found to differ significantly between the two sexes consistently across production contexts: fundamental frequency (f_0), the first formant (F_1) bandwidth, and cepstral peak prominence (CPP). Many other acoustic measurements were found to differ significantly between men and women; however, these measurements were indicated only in certain contexts.

The second question addressed in the current study examined the perception of gender by listeners and the important characteristics of the voice that indicate the gender of the speaker. The mean f_0 was found to be the primary indicator for all contexts. The following predictors varied between the different contexts and the method of acoustic analysis: combined with f_0, voice turbulence index (VTI) and soft phonation index (SPI), as well as H1-A3, were found to be significant predictors for gender perception when listening to a sustained vowel; for both carrier phrases of the measurements taken from the isolated vowels, H1-H2 and f_0 were found to be significant indicators of gender. Finally the combination of the measurements taken from the entirety of the carrier phrases, f_0 and CPP, was found to be a significant predictor of listeners' judgements of speaker gender. Further slight variation between carrier phrases existed with the addition of certain measurements to the models (i.e. percentage of jitter and f_0 standard deviation across phrase).

5.2 WHAT ACOUSTIC MEASUREMENTS OF VOICE QUALITY CUE LISTENERS TO THE PERCEPTION OF ONE SEX OR THE OTHER?

5.2.1 Fundamental frequency

In previous studies vocal pitch has been described as "...a strong gender marker" (p.520, Holmberg, Oates, Dacakis, and Grant, 2010). The current study supports previous findings in that fundamental frequency (f_0) and measurements of its variability (i.e. range and standard deviation) were found to be significantly different between male and female voices, and to be a significant predictor of listener judgement of gender. The means of f_0 for the three contexts – i.e. sustained vowel, carrier phrases with *had*, and carrier phrases with *hodd* - were 194.29 Hz, 190.39 Hz and 191.55 Hz for females, respectively, and 128.60 Hz, 128.86 Hz and 128.59 Hz for males, respectively. These values fall within or near to the expected range for adult males (107-146 Hz) and adult females (196-224 Hz) (Casper, Colton and Leonard, 2006).

Exceptions did exist for participants mean f_0 outside of the expected range. Two male participants produced a high mean f_0, when compared to the other male participants, and close to the female participants' mean f_0 – i.e. approximately 190 Hz. M48 achieved a mean f_0 of 186.95 Hz, followed by M39 185.78 Hz for production of the sustained vowel. Although female mean f_0 was within or close to the expected female range, F9 and F27 were two female participants who achieved the lowest mean f_0 for sustained vowel production for the female group (157 Hz and 144.71 Hz, respectively). These values place some participants into the gender ambiguous pitch range as described by Mordaunt (2006): a voice at 150 Hz to 185 Hz can still be perceived as female. These results are indicative of how variable the phonation pitch range is for men and women.

The f_0 range of the transgender/puberphonic voice group is within the ambiguous pitch range (150 Hz to 185 Hz) and some within the female range 196-224 Hz (Casper, Colton and Leonard, 2006). The f_0 in the sustained production of the vowel ranged from 248.73 Hz for participant T59, the individual with puberphonia, to 154.21 Hz for participant T62. For the carrier phrases, the mean f_0 of the entire phrase ranged from

239.47 Hz for T59 to 149.68 Hz for T63. The f_0 of the individual with puberphonia was indicated as producing a mean f_0 well above the other transgender group members for two of the contexts: sustained vowel and the isolated /ɒ/ from *hodd*. Although participants were within ambiguous pitch range and sometimes within the female pitch range, they were not always identified as the female gender.

As illustrated in figures 4 to 6, the scatterplots of mean fundamental frequency and perception of gender, although f_0 is indicated as a primary difference between male and female voices some voices deviated from the group trend. Voice cues related to intonation or quality are shown to predict the perception of gender by listeners.

5.2.2 Intonation

The parameters measured for evaluating intonation were mean f_0, standard deviation of f_0, highest and lowest frequency, as well as range and standard deviation of semitones. These measurements were found to be significantly different between men and women in a previous investigation by Skodda, Visser, and Schlegel (2009). All the measurements described in the current study were taken from the two carrier phrases containing *had* and *hodd*. No significant difference was found for the semitone range and standard deviation between male and female voices, but a significant difference was found for f_0, including all related measurements (i.e. highest, lowest, mean and standard deviation). In the current study, the female mean range for reading a phrase in semitones was 14.03-14.13 with a standard deviation in semitones of 3.27–3.15, and male mean range was 11.46-12.5 with SD in semitones of 2.93-2.77. The normative standard deviation for reading in adult females ranges from 2.5 to 4.7 semitones and is approximately 3.3 semitones for adult males (Casper, Colton and Leonard, 2006). The female participants of this study were within the normal range as well as the male voices. Despite the smaller mean value of semitone range and standard deviation for men when compared with women, no significant difference was found.

Participant T63 was indicated as female by more listeners in the carrier phrase with *hodd* than T59, even though T59 had a higher f_0 and greater values of the different measurements used to examine intonation. Participant T62, who had the broadest range of frequency as indicated by maximum and minimum frequency, was indicated as female in this same context by only two listeners. Maximum lowest and highest frequency may not be accurate in the measure of the range used in speech intonation due to possible artifact in the environment, as recordings were taken in soft but not sound proofed areas. Even though the mean measurements associated with intonation in the female voice are found to be higher than in the male voice, listeners still indicated some participants as female who did not follow this trend, as seen with listener response to participant T63 in the carrier phrase with *hodd*. This result and the lack of significant difference between semitone range and standard deviation found between men and women is unexpected.

In the previous study by Skodda, Visser, and Schlegel (2009), range in semitones was accepted as a measurement addressing intonation and found to differ between male and female speakers; however, this study was conducted in German opposed to the current study in English. The speaker`s range as measured by semitones may be subject to language or sociolinguistic differences. These findings may support the concept that primarily in the English language intonation is standard across gender (Mordaunt, 2006).

A possible cause for the discrepancy observed in the non-significant findings for differences in intonation in this study may be the difference in intonation analysis. Wolfe et al. (1990) also found significant differences between men and women for their measurements of intonation (i.e. mean extent of upward and downward shifts of frequency in semitones; and percentage of upward, downward, and level semitone shifts). The current study limited intonation analysis to range of semitones and did not examine the percentage of upward or downward shifts.

Wolfe et al. (1990) used conversational responses to questions about home and work, where the current study required participants to read a standard text. Gelder and Schofield (2000) found no significant differences between male and female voices from read passages for their measurement of intonation. One possible explanation these authors proposed to explain this discrepancy from other studies was that the recorded material was not spontaneous speech but rather readings taken from the Rainbow Passage. Similarly in the current study, the recorded material was read and not taken from spontaneous, conversational responses. Eliciting a reading task versus a spontaneous speech sample may have influenced the variance in the fundamental frequency.

5.2.3 Voice quality

The investigation of acoustic measurements associated with perception of gender was conducted using two main contexts: sustained vowel /ɒ/ and carrier phrases. The two carrier phrases differed in the vowel included in the /hVd/ context, i.e. /æ/ from *had* and /ɒ/ from *hodd*. Along with measurements associated with voice quality taken from the isolated vowels from these carrier phrases, ceptral peak prominence (CPP) and cepstral prominence-smoothed (CPPs) were taken from the entire phrase. Many measurements were found to be significantly different between male and female voices, but only two measurements related to voice quality were found to differ in every context: first formant (F_1) bandwidth and CPP.

The current study found there was a significant difference between the sexes in F_1 bandwidth with females producing a larger F_1 bandwidth than males, following the same trend as the data from Hanson and Chuang (1999); however, a number of outliers were found for each group. For the male participants the number of outliers for each recorded context varied from two to three participants.

The participants from the transgender/puberphonic voice group with the highest and lowest value of F_1 bandwidth varied between the different vowels (i.e. /ɒ/ and /æ/) and context (phrase and sustained productions). Participant T61 had the lowest recorded F_1 bandwidth for the vowel /ɒ/ in both contexts. Participant T59 had the highest recorded F_1 bandwidth for vowels isolated from the two different carrier phrases. For the reason that F_1 bandwidth has been associated with breathiness (Hanson and Chuang, 1999), this may indicate that the individuals in the transgender group may be attempting to speak with higher mean f_0 as a result of laryngeal tension combined with incomplete vocal fold closure. Rowan-Gorham and Morris (2006) concluded in their study of the aerodynamic measurements of MtF transgender voice that some transgender participants were using a combination of laryngeal tension and incomplete vocal fold closure. High minimum airflow rate was recorded possibly implicating that MtF participants closed their vocal folds quickly but incompletely when phonating. Increased airflow during phonation results an increase in breathiness in the voice.

When examining the results of the multiple regression analysis, F_1 bandwidth was not considered to be significant in a predictive model of listener perception of gender. Although this measurement was found to be significantly different between men and women, listeners do not appear to attend or perceive this as a cue when judging gender. This may indicate that the mid-frequency noise, in the region of F_1, is not of great importance when listeners perceive gender or that the measurement of F_1 bandwidth is insufficient in capturing the details that listeners are attending to in the voice.

Another measurement related to glottal configuration, H1-H2, was found to be part of a model that included statistically significant predictors in gender perception. The measurement of the H1-H2 amplitude relation has been shown to be significantly correlated to abduction (open) quotient in simultaneous observations of airflow and acoustic spectra for female speakers (Holmberg et al, 1995 as cited by Hanson, 1997). Although found to be part of a predictive model along with f_0 in gender perception for the carrier phrases, a significant difference was found between male and female voices only in the sustained vowel production. The mean of H1-H2 amplitude relations in the sustained vowel context for women was 4.49 dB (± 4.30) and 2.07 (± 3.26) for men – i.e. a difference of 2.4 dB between women and men. For the isolated vowels, the H1-H2 mean for female voices was larger than males but not significantly different. Previous research utilizing the H1-H2 amplitude relation shows larger differences between male and female voices than the results presented here. For example Klatt and Klatt (1990) found a 5.7 dB difference between female and male voices. The current results – i.e. difference between female and male voices is 2.4 dB - are closer to observations made by Hanson (1996), and Hanson and Chuang (1999). These authors found a difference of about 3 dB between the grand average of H1-H2 harmonic relations produced by the sexes. The current results follow the general trend in that female speakers on average have larger relative amplitudes of the first harmonic. This suggests that females have larger open quotients (Hanson and Chuang, 1999).

The grand average of H1-H2 relations for the transgender group was 5.68 dB indicating a large difference from the biological female speakers (1.7dB) and an even larger difference from the male voice group (i.e. 4.1 dB). The highest value of relative amplitude of H1 within the transgender group was recorded from participant T60. This participant was relatively early in her transition and had recently begun to alter her voice. Laryngeal tension used to increase her pitch may have influenced her open quotient reflected in the consistent measure of a large H1-H2.

The amplitude of the first harmonic (H1) indicated by comparison of H1-H2 was found by Hillenbrand and Houde (1996) to show a moderate correlation (0.52) to perception of breathiness in connected speech, as opposed to a moderately strong correlation (0.70) in sustained vowel production. This was attributed to the strong effect variations in phonetic context and laryngeal posture has on H1-H2 amplitude relations. Shrivastav and Sapienza (2003) found a moderate correlation (0.55) between H1-H2 and perception of breathiness in a sustained vowel /ɒ/. Despite a stronger correlation, as found in previous studies, H1-H2 relation was found to be relevant to listeners only in connected speech.

Another measure of amplitude relation, H1-A1, was found to be significantly different between men and women for certain contexts, but not found to be part of a predicative model of gender perception. This is not wholly unexpected as H1-A1 is related back to the measurement of F_1 bandwidth (Hanson, 1997). As with F_1 bandwidth and H1-H2 amplitude relations, this may indicate that listeners attend or process different acoustic cues and may place more or less focus on these cues depending on the situation – i.e. connected speech versus sustained vowel. Although the male voice and the female voice may significantly differ between certain vocal parameters, according to these three acoustic characteristics, listeners may not always incorporate acoustic cues, or may not incorporate all available acoustic cues, into their perception of gender.

The second measurement consistently identified across contexts as significantly different between male and female voices was cepstral peak prominence (CPP). The amplitude of the cepstral peak, as measured by CPP, reflects the periodicity of the acoustic signal. The mean CPP for the sustained vowel production was found to be significantly smaller in female voices than male voices. Smaller CPP has been associated with a greater aperiodic signal, possibly related to breathiness (Hillenbrand and Houde, 1996). A significant difference was found between male and female voices in the carrier phrase conditions; however, the male voice production of CPP and CPPs was significantly smaller than the female voices in these contexts. This may indicate that the male voice had an increase in aperiodicity in connected speech. Awan and Roy (2005) suggested that a female voice with an increase in breathy quality can more commonly be mistaken as normal if the signal has a relatively strong periodicity despite the excessive airflow during phonation. The measurement of CPP may not necessarily have recorded the additive noise of the turbulent airflow in the voice in connected speech, solely the aperiodicity of the signal.

Interestingly, with the additional step of smoothing the CPP measurement, the CPPs was not found to be significantly different in the carrier phrases for men and women; however, it remained significantly different in the sustained vowel production. Methods differ in the smoothing window size for measuring CPPs in sustained vowel production as opposed to connected speech. Larger smoothing windows across time and relatively smaller smoothing windows across cepstral domain work best for sustained vowels as opposed to smaller time-smoothing windows needed to accommodate for the variation of f_0 in connected speech (Hillenbrand and Houde, 1996). The measurements of CPP and CPPs have been related to detectable differences between breathy and non-breathy voices by Hillenbrand and Houde (1996), but not necessarily to differences in male and female voices. Male and female voices had similar values for the mean of CPPs for the carrier phrase production, possibly indicating equal aperiodicity of the signal for these speakers.

The mean CPP value for female participants sampled from sustained vowel production was 18.8 (± 3.01) and the mean for male participants was 20.52 (± 2.89). All of the participants in the transgender group fell within the female range for this context, exhibiting CPP values in the upper range of the female participants. The highest value for CPP was found for participant T62, and the lowest value for CPP was found for participant T63. For CPPs the lowest value was attributed again from T63. The highest value was from T60, followed by T62. For the two carrier phrase conditions, the smallest value for CPP was found in participant T59 and highest in T62.

Several other acoustic measurements were shown to be significantly different between male and female voices between the two carrier phrases. The mean value of shimmer in the male participants was found to be significantly greater than the female group mean for both isolated vowels. Awan and Roy (2005) found that shimmer was correlated to a rough or hoarse classification of voice. The authors also found that none of the females with hoarse voice, in their investigation into listener classification of female disordered voices, were identified as having a normal voice, as opposed to female breathy voices which were often classified as normal. The current study found that there is a significant difference between the sexes in degree of amplitude perturbation but only for samples taken from connected speech. The percentage of jitter was found to be significantly different only for the isolated /ɒ/ vowel taken from the carrier phrase. Although not found to be significantly different, the mean value for percentage of jitter taken from the sustained vowel was greater in the female group (1.06%>.71%). Increased breathiness of the voice is correlated to a larger percentage of pitch perturbation in the voice (Shrivastav and Sapienza, 2003). Conversely, the male participants exhibited greater percentage of jitter in the voice as measured in the isolated vowels sampled from the carrier phrases.

The method of analysis for determining the values of percentage of jitter and shimmer from the isolated vowels may impact the results. The CSL program required that the signals be resampled, as none of the recordings were taken by the program itself. Using the PRAAT program, all recordings for analysis were resampled at 25 KHz. Both male and female means were found to be outside the normative values of both percentage of jitter and shimmer, which may indicate that the sampling method used to measure these parameters was less than ideal and may have altered some acoustic parameters during the resampling. The program analyzes signals by obtaining the period of each cycle of vibration, subtracting it from the previous or succeeding period, averaging the differences, dividing by the average period and multiplying by 100 (Colton, Casper and Leonard, 2006). With a short signal sample, as in the isolated vowels taken from the carrier phrases, the results may not be completely reliable. The impact of using such short samples may have also affected the validity of the other measurements from CSL, taken from the isolated vowels (i.e. SPI from /ɒ/ in *hodd*).

It is possible that the percentage of jitter and shimmer in the voice may impact the measurement of CPP. The CPP measurement does not just examine the noise that is due to an aspiration source but to the noise that results from a combination of factors, i.e. shimmer and jitter (Hanson, 1997). This may be a possible explanation for the difference in findings for CPP and CPPs measurements for male and female voices in connect speech samples.

The fact that participant T59, in the production of both carrier phrases had a high F_1 bandwidth and a low CPP measurement may indicate this participant was exhibiting the voice characteristics related to an aperiodic signal. This would appear logical in that the individual has a diagnosed voice disorder, i.e. puberphonia. The variability of the different measurements for each participant indicates the wide variation which exists among speakers, and due to the low sample number may not be generalizeable.

5.3 DO ACOUSTIC MEASUREMENTS OF VOICE QUALITY CORRELATE TO THE PERCEPTION OF A SPEAKER'S MASCULINITY OR FEMININITY?

5.3.1 Two voices from one participant

To increase the femininity of the voice quality of a speaker, often a higher pitch and increased vocal breathiness is attempted (Gorham-Rowan and Morris, 2006). The results of the current study support the previous findings for the relationship between increasing f_0 and perception of a voice as increasingly female; however, as previously mentioned data presented in figures 4 to 6 of rated gender versus mean $f_{0,}$ some voices deviated from the group trend. Holmberg, Oates, Dacakis, and Grant (2010) in their research on the male-to-female transsexual voice found that although pitch was indicated as a primary indicator in gender perception, other factors correlated to perceptual ratings. The results of the current study indicate the importance of voice quality and suprasegmental cues above and beyond pitch in the listener's perception of gender.

A closer examination of the data taken from participants T59 and T60 may help illustrate the relative importance of pitch and gender perception. Both of these participants were recorded as biological male voices and as transgender voices. Participant T59 is not a transgender individual; however, his unique pitch and voice quality due to puberphonia provided an opportunity to examine the impact a change in mean f_0, along with its related measurements, has in the perception of the same voice. Participant T60 also provided such an opportunity as she was still transitioning into her gender appropriate voice and was still comfortable using her male voice.

As seen in table 14, the participant with puberphonia produced a very effective feminine voice when using his puberphonic voice. When he produced his male voice the mean f_0 for the sustained vowel was still within the gender ambiguous range which may account for the low average score for masculinity. Although he was identified as a male while using his male voice for the carrier phrases, his score for masculinity was low. In these conditions his mean f_0 was within or just above the upper limit of the average male f_0. Gorham-Rowan and Morris (2006) concluded that transsexuals in their study produced an effective feminine voice through increased laryngeal tension in combination with incomplete vocal fold closure. The falsetto register, as produced by participant T59/M44, is characterized by a strong longitudinal tension in the vocalis ligament (Sundberg and Hogset, 2001). As well, when the voices of MtF transsexuals were compared with biological male voice production, MtF transsexuals had higher alternating glottal airflow and maximum flow declination rate. This may indicate a more abrupt shut off of glottal airflow due to the increased tension. It has been found that MtF transgender speakers after spending years living full time as women may retain certain characteristics of female voice productions when speaking in their male voices (Gorham-Rowan and Morris, 2006). Since participant T59/M44 has increased laryngeal tension it is possible that the vocal-fold closing is more abrupt for him, increasing maximum flow declination rate. This laryngeal tension may carry over into his lowered pitch voice, as labeled M44, and result in an abrupt vocal fold adduction influencing his airflow rate.

The opposite may be true for participant T60/M46. Her male voice was very effective but in her female voice she was consistently identified as male. Her gender score, however, does reveal that some of the characteristics she is beginning to apply may be influencing her listener’s perception of her voice. Due to the early progression into her transition, T60 may still be relying on excessive laryngeal tension to achieve her higher pitch. Higher pitch as achieved by excessive laryngeal tension may result in perception of a strained vocal quality (Gorham-Rowan and Morris, 2006). Possibly this strained vocal quality, combined with other speech parameters such as intonation, contributed to the listener’s perception of T60’s gender. She was identified with a higher femininity rating when listener’s perceived the sustained vowel production as opposed to the two carrier phrases.

Listeners were generally more reliable for inter- and intra-rater reliability for the perception of the carrier phrases as opposed to the sustained vowel segments. This may indicate the amount of information that was available to the listener for each stimulus differed between the two conditions: short segmented sustained vowel versus a carrier phrase. Connected speech includes information unavailable in sustained vowel production to the listener, which has been described in the literature as influencing gender perception such as rate and articulation (Adler, 2006). The listener was restricted to a highly unnatural sounding voice segment when judging the sustained vowel. In this context, different cues may have had greater influence and choice of what components to base one’s judgment on may be individual. The decrease in both inter-rater reliability and percent identification of gender in the sustained vowel from the carrier phrases is also likely due to availability of additional prosodic information in the carrier phrases. Segments were often described by listeners as sounding computerized, possibly due to the absence of a natural onset and offset for most of the segments. The best model for prediction of gender perception differed between the sustained vowel context and the carrier phrases, possibly due to the different data and duration available to the listener.

5.3.2 Perception of connected speech

The combination of f_0 and H1-H2 amplitude relations were found to be significant indicators correlated to perception of gender in both carrier phrases. This model included those measurements taken from the isolated vowels of the words *hodd* and *had* in the two different phrases. Changes in the relative amplitudes of H1 and H2 have been shown to reflect changes in the open quotient of the glottis (Hanson and Chuang, 1999). Differences in the open quotient are reflected in the spectrum as changes to the low frequencies in the voice, i.e. H1 and H2. The female voice has been found to have a larger relative first-harmonic amplitude on average when compared to male voices (Klatt and Klatt, 1990). The association of gender rating and the relative amplitude of H1 supports the previous findings of Klatt and Klatt (1990). In the perceptual data examined by Klatt and Klatt (1990) the amplitude of H1 relative to H2 was closely tied to the perception of breathiness as sampled from an isolated vowel taken from a sentence; more so than any other measurement explored in their study, which included H1-A3 amplitude relations. These authors concluded that as the relative first-harmonic amplitude (H1) is an acoustic correlate of breathiness, and that females are more breathy than males.

Percentage of jitter was also found to be a significant predictor of perception of gender when combined with f_0 and H1-H2 amplitude relations but only for the isolated vowel /ɒ/ from the carrier phrase with *hodd*. Percent jitter was not found to be significantly correlated to gender perception in the vowel /ɒ/ taken from the sustained vowel production of speakers. Shrivastav and Sapienza (2003) found a strong correlation of percent jitter to breathiness as measured from the vowel /ɒ/. These authors analyzed samples 500-ms in duration for their study. Although the listeners heard short segments of the original sustained vowel sample, the original production was analyzed and was approximately 10 seconds in length, compared to the analyzed isolated vowel from connected speech approximately 400-ms. Possibly the discrepancy between the two contexts resulted from the influence of analyzing the same vowel from productions of differing duration – i.e. connected speech versus sustained vowel.

Isolated vowel measurements may be impacted by the transitions between the vowel and adjacent consonants in connected speech despite the effort to isolate the vowel from the influence of coarticulation. In perception of the carrier phrases, listeners had a greater amount of time and information upon which to base their perception of gender. As indicated by Adler (2006), many parameters are involved in the overall perception of connected speech as one gender or another: intonation, stress, rate, articulation and volume of speech may have influenced listener judgement. The lack of control over these components may have impacted the results. The measures CPP and CPPs were used to explore the voice quality as analyzed from connected speech.

CPP was indicated as a good predictor of gender perception when combined with f_0. Previously in the literature, CPP has been correlated to perception of breathiness for both sustained vowel production and connected speech. Hillenbrand and Houde (1996) found a strong negative correlation between perception of breathiness and CPP for normal speakers. Shrivastav and Sapienza (2003) also found a strong negative correlation for CPP and perception of breathiness for women with dysphonia. In the current study, CPP is indicated as having a negative correlation with gender rating in connected speech, which suggests that rating of masculinity (arbitrarily designated as an increasing number (0 to 100), as opposed to a decreasing number for femininity (0 to -100)) by listeners increases as CPP decreases. A smaller CPP is indicated in the literature as correlating to an increase in breathiness perception. This disputes many of the assumptions brought forth by studies, including the current one, regarding the association between the perception of breathiness and femininity of the voice (Van Borsel, Janssens, and De Bodt, 2009).

The CPP and CPPs measurements taken from the sustained vowel production followed the prediction of the current study that a measurement associated with breathiness perception would correlate to perception of gender; specifically, the more feminine voice would have a smaller CPP which is associated with greater perception of breathiness. Since the scale designated to measure femininity is based on a negative scale - i.e. 0 (least feminine) to -100 (most feminine), a smaller gender score indicates a higher feminine rating. Hillenbrand, Cleveland, and Erickson (1994) found that CPP measurements correlated negatively to perception of breathiness in normal speakers when sampled from a sustained vowel (approximately 3 seconds). Hillenbrand and Houde (1996) also found a negative correlation with CPP and CPPs to breathiness in dysphonic speakers in sustained vowel and connected speech samples. In the current study because CPP and CPPs is found to have a positive correlation to perception of gender in connected speech samples, increased breathiness is implied as a cue to listener perception of increased masculinity (i.e. smaller CPP, greater perception of breathiness and a more positive or masculine rating).

One explanation to this unexpected result may be related to a complicating factor identified by Hanson and Chuang (1999) in their analysis of male voices. These authors found a second pulse during a glottal period of some of the male participants. Extra pulses interfere with the periodicity of the waveforms and may hamper the confidence of listeners when judging noise in the voice. Considering that CPP is a measurement of the harmonic organization of the signal, this second excitation pulse may impact listener perception of periodicity, noise and possibly gender. Hanson and Chuang suggested that this phenomenon may be dependent on the vowel sampled – i.e. extra pulses were more often identified in vowels with a wide F_3 bandwidth. They also offered that a large amount of surface tension of the vocal folds may influence the presence or absence of this extra pulse. This phenomenon was not observed for the female participants or for all the male participants in their study, and was not examined in the current research.

The inclusion of CPP into a predictive model for the samples of connected speech with a negative correlation with gender perception reflects the impact the combination voice quality and level of f_0 may have in gender perception. Hillenbrand and Houde (1996) used CPP to measure the connected speech of individuals with dysphonia and found a significant correlation between the perception of breathiness and CPP. The subjects of their study were chosen to represent a range of breathy voices. This perception of aperiodic signal combined with a low pitch may cue the listener to perceive the speaker as hoarse. Hoarseness is described as reflecting aperiodic vibration of the vocal folds, a disorder many of the subjects from the study by Hillenbrand and Houde (1996) experienced (Colton, Casper, and Leonard, 2006). Hoarseness has been implicated as a significant indicator of a normal male voice (Singh and Murray, 1978). Evidence that perception of hoarseness of the voice may play a role as a cue to gender of the voice is implicated in the significant difference found between male and female voices in the measurement of shimmer from the carrier phrases and not the sustained vowel production. The relationship of f_0 and CPP measurement may support this particular assumption.

When all of the measurements, isolated from the vowel and sampled from the entire phrase, were included into a multiple linear regression model, significant predictors were found to vary slightly between contexts. A combination of f_0 and H1-H2 amplitude relation was found to be significant predictor along with the two different cepstral measurements: CPP for the phrase with *hodd* and CPPs for the phrase with *had*. It appears that the perception of a feminine voice from connected speech requires cues such as an increased f_0 with an increased perception of breathiness primarily represented as aspiration noise affecting the lower portion of the spectrum – i.e. difference between H1-H2 - as well as a more periodic signal – i.e. associated with larger CPP or CPPs value. As suggested earlier a female voice may be perceived as normal, although breathy, if the signal has a relatively strong periodicity despite the excessive airflow during phonation (Awan and Roy, 2005).

5.3.3 Perception of sustained vowel

Many voice quality measures were indicated as significant predictors of gender perception when combined with f_0 for the sustained vowel: voice turbulence index (VTI), H1-A3 amplitude relations and soft phonation index (SPI). The predictor second to f_0, VTI, is a measure of the high frequency energy in the range of 2800-5800 Hz (high) relative to harmonic energy in the range of 70-4500 Hz (low to mid). Unlike NHR, this is not a general evaluation of noise present in the analyzed signal. SPI could be considered the opposite of VTI, in that it is an average of the ratio between the lower-frequency harmonic energy in the range of 70-1600Hz (low to mid) to the higher-frequency harmonic energy in the range 1600-4500 Hz (mid to high). A positive correlation was found between the perception of gender and VTI, and a negative correlation between the SPI and perception of gender. This indicates that if the speaker has a high pitch, a small VTI (low relative energy level of high frequency noise), larger SPI (higher amounts of harmonic in low to mid frequency energy) the person may be perceived as more feminine.

The final measurement included in the predictive model for gender perception of sustained vowel production was the H1-A3 amplitude relations. This is a measure of source spectral tilt (Hanson and Chuang, 1999). As the difference between the amplitudes of H1 and A3 increased, i.e. spectral tilt increased, the voice was perceived as more feminine. Increased spectral tilt or slope with lesser high-frequency harmonic energy may be the result of increased glottal airflow (Holmberg et al, 2010). The significance of H1-A3 measurement is congruent with the relationships described for VTI and SPI with gender: greater harmonic energy in the low frequency with a drop in energy in high frequency energy– i.e. steep spectral tilt. This supports the results of previous literature which found that an increased spectral tilt was indicative of a female voice, as opposed to a male voice (Hanson, 1996; Hanson and Chuang, 1999; Klatt and Klatt, 1990).

5.4 CONCLUSION

Many apparent contradictions exist between the acoustic measurement data collected corresponding to actual gender and the data for perceived gender. For example F_1 bandwidth was shown to be significantly different between men and women, but was not found to be significant in a predictive model of listener perception of gender. Nonetheless, fundamental frequency was consistently shown to discriminate actual speaker gender, and was also a good correlate to perceived gender. A complex combination of vocal parameters is likely involved in the gender perception of a speaker. Further investigation is needed to identify the exact role and context dependency of voice quality in gender identification.

Measures that should be examined more closely through future research as related to voice quality and as possible indicators of gender are the F_1 bandwidth, cepstral peak prominence measurements (CPP and CPPs), H1-H2 and H1-A3 amplitude relations, and the parameters measuring frequency noise in specific regions of the signal, i.e. voice turbulence and soft phonation index. Despite significant differences found between male and female voices, these acoustic characteristics may not be as important to the listener making the judgement of female gender as are a greater spectral tilt and a high level of mean f_0. However, this conclusion and the means of measuring the acoustic characteristics of the voice appear to depend on the type of acoustic signal the listener hears - i.e. connected speech versus sustained vowels. When listeners hear a voice in connected speech, they may rely on a high level of f_0 and periodicity of the signal as cues to the speaker's femininity. Conversely, a listener may associate an aperiodic signal, as indicated by a lower CPP, with a feminine voice quality when hearing a sustained vowel. As well, a listener may be attending to more complex prosodic cues such as intonation when listening to connected speech production.

Future research should include more rigorous intonation measurements to explore the differences between the acoustic characteristics of the voice and speech between men and women, which listeners attend to and process in different speech/voice samples. More natural contexts may help to inform how the acoustic measurements explored in this study actually cue and influence gender perception in real conversation.

Studies including a greater number and variety of transgender participants are needed to make any general conclusions of the information gathered here. Research in the future could consider including self-evaluation and satisfaction into data collection to increase the ecological validity of the information. Applying objective measures that may indicate the likelihood of the target response for gender perception is only half of the story: including the transgender client’s level of satisfaction and confidence in using her voice is of equal importance, especially in daily living.

REFERENCES

Adler, R.K. (2006). Chapter 1: Transgender/transgender: An understanding. In Adler, R.K., Hirsch, S., and Mordaunt, M. (Eds.) *Voice and Communication Therapy for the Transgender/Transgender client: A Comprehensive Clinical Guide* (1-14). San Diego, CA: Plural publishing.

American Psychiatric Association (2011) "DSM-5 Development; Proposed Revisions, 302.3 Transvestic Fetishism," http://www.dsm5.org/ProposedRevision/Pages/proposedrevision.aspx?rid=189#

American Psychiatric Association (2011A) "DSM-5 Development; Proposed Revisions, P 01 Gender Dysphoria in Adolescents or Adults," http://www.dsm5.org/ProposedRevision/Pages/proposedrevision.aspx?rid=482

Andrews, M.L. & Charles, P.S. (1996). Gender presentation: Perceptual and acoustical analyses of voice. *Journal of Voice, 11(3),* 307-313.

Avery, J. & Liss, J.M. (1996). Acoustic characteristics of less-masculine-sounding male speech. *The Journal of the Acoustical Society of America, 99(6),* 3738-3748.

Boersma, P. (2001). PRAAT, a system for doing phonetics by computer. *Glot Int, 5*, 341-345.

Boersma, P., & Weenink, D. (2001). *PRAAT, system for doing phonetics by computer (version 4.3.14) computer program*. Institute of Phonetic Sciences: Amsterdam, Netherlands. Retrieved from http://www.praat.org.

Chang, A. (2011). Cohen's and Fleiss's Kappa program. *Stattools.* Retrieved from http://www.stattools.net/CohenKappa_Pgm.php.

Childers, D.G. & Wu, K. (1991). Gender recognition from speech. Part II: Fine analysis. *The Journal of Acoustical Society of America, 90(4),* 1841-1856.

Coleman, R.O. (1971). Male and female voice quality and its relationship to vowel formant frequencies. *The Journal of Speech and Hearing Research, 14(3),* 565-600.

Coleman, R.O. (1976). A comparison of the contributions of two voice quality characteristics to the perception of maleness and femaleness in the voice. *Journal of Speech and Hearing Research, 19,* 168-180.

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

Colton, R.H., Casper, J.K. & Leonard, R. (2006). Chapter 2: Differential diagnosis of voice problems. In Colton, R.H., Casper, J.K. & Leonard, R. *Understanding*

Voice Problems: A physiological perspective for diagnosis and treatment (12-63). Lippincott Williams & Wilkins: Philadelphia, PA.

Colton, R.H., Casper, J.K. & Leonard, R. (2006). Chapter 4:Phonotrauma: Its effects on phonatory physiology. In Colton, R.H., Casper, J.K. & Leonard, R. *Understanding Voice Problems: A physiological perspective for diagnosis and treatment* (76-117). Lippincott Williams & Wilkins: Philadelphia, PA.

Colton, R.H., Casper, J.K. & Leonard, R. (2006). Chapter 14: Some normative data on the voice. In Colton, R.H., Casper, J.K. & Leonard, R. *Understanding Voice Problems: A physiological perspective for diagnosis and treatment* (410-420). Lippincott Williams & Wilkins: Philadelphia, PA.

Dacakis, G. (2002). The role of voice therapy in male-to-female transgenders. *Current Opinion in Otolaryngology and Head and Neck Surgery, 10,* 173-177.

Eskenazi, L., Childers, D.G., & Hicks, D.M. (1990). *Acoustic correlates of vocal quality. Journal of Speech and Hearing Research, 33,* 298-306.

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

Gender Education & Advocacy, In. (2001). Basic tips for health care and social service providers for working with transgendered people. *Gender.org*^{tm} . Retrieved from: http://www.transgendered-soul.com/TransitionPDF/Basictipsforhealthcare2.pdf.

Goldwave. (2010). *Goldwave v5.58*. Goldwave, Inc.: St. John's, ND. Retrieved from http://www.goldwave.com/.

Gunzburger, D. (1995). Acoustic implications of the transsexual voice. *Archives of Sexual Behaviour, 24(3),* 339-348.

Hancock, A. Krissinger, J., & Owen, K. (2011). Voice perceptions and quality of life of transgender people. *Journal of Voice, 25(5),* 553-558.

Hanson, H.M. (1997). Glottal characteristics of female speakers: Acoustic correlates. *Journal of Acoustical Society of America, 101(1)*, 466-481.

Hanson, H.M. & Chuang, E.S. (1999). Glottal characteristics of male speakers: Acoustic Correlates and comparisons with female data. J*ournal of Acoustical Society of America, 106(2),* 1064-1077.

Hillenbrand, J., Cleveland, R.A., & Erickson, R.L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research, 37*, 769-778.

Hillenbrand, J. & Houde, R.A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research, 39,* 311-321.

Hixon, T. Weismer, G., & Holt, J. (2008). Chapter 6: Acoustics. In Hixon, T. Weismer, G., & Holt, J *Preclinical Speech Science: Anatomy, Physiology, Acoustics, Perception* (317-356). San Diego, CA: Plural Publishing.

Hixon, T. Weismer, G., & Holt, J. (2008). Chapter 7: Acoustic theory of vowel production. In Hixon, T. Weismer, G., & Holt, J *Preclinical Speech Science: Anatomy, Physiology, Acoustics, Perception* (357-400). San Diego, CA: Plural Publishing.

Holmberg, E., Oates, J., Dacakis, G., & Grant, C. (2010). Phonetograms, aerodynamic measurements, self-evaluations, and auditory perceptual ratings of male-to-female transsexual voice. *Journal of Voice, 24 (5)*, 511-522.

Kay Elemetrics Corp. (2004). *Multi-dimensional voice program model 5105*. Kay Elemetrics: Lincoln Park, USA.

Kay Elemetrics Corp. (2004). *Real time pitch program*. Kay Elemetrics: Lincoln Park, USA.

Klatt, D.H. & Klatt, L.C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of Acoustical Society of America, 87(2),* 820-856.

Lass, N.J., Hughes, K.R., Bowyer, M.D., Waters, L.T., & Bourne, V.T. (1976). Speaker sex identification from voiced, whispered, and filtered isolated vowels. *Journal of Acoustical Society of America, 59(3),* 675-679.

Lawerence, A. A. (2001) Vocal feminization surgery: A critical overview. Transgender women's resources. Retrieved June 3, 2010 from http://www.annelawrence.com/twr/.

Mathworks. (2011). *MATLAB*. Mathworks, Inc. Retrieved from http://www.mathworks.com.au/products/matlab/.

McNeill, E.J.M. (2006). Management of the transgender voice. *The Journal of Laryngology and Otology, 120,* 521-523.

McNeill, E.J.M., Wilson, J.A., Clark, S., & Deakin, J. (2008). Perception of voice in the transgender client. *Journal of Voice, 22(6)*, 727-733.

Mendoza, E., Valencia, N., Munoz, J., & Trujillo, H. (1996). Differences in voice quality between men and women: Use of the long-term average spectrum (LTAS). *Journal of the Voice, 10(1),* 59-66.

Mordaunt, M. (2006). Chapter 10: Pitch and intonation. In Adler, R.K., Hirsch, S., and Mordaunt, M. (Eds.) *Voice and Communication Therapy for the Transgender/Transgender Client: A Comprehensive Clinical Guide* (168-208). San Diego, CA: Plural publishing.

Murray, T. & Singh, S. (1980). Multidimensional analysis of male and female voices. *Journal of Acoustical Society of America, 68(5),* 1294-1300.

Nittrouer, S., McGowan, R.S., Milenkovic, & P.H., Beehler, D. (1990). Acoustic measurements of men's and women's voices: A study of context effects and covariation. *Journal of Speech and Hearing Research, 33,* 761-775.

Perry, T.L. (2001). The acoustic bases for gender identification from children's voices. *The Journal of Acoustical Society of America, 109(6),* 2988-2998.

Pisanski, K. & Rendall, D. (2011). The prioritization of voice fundamental frequency of formants in listeners' assessments of speaker size, masculinity, and attractiveness. *Journal of the Acoustical Society of America, 129(4),* 2201-2212.

Portney, L.G. & Watkins, M.P. (2009). Chapter 20: Comparing more than two means: Analysis of variance. In Cohen, M (Ed.) *Foundations of clinical research: Applications to practice 3rd ed.* (450-478). Pearson Education, Inc.: New Jersey, USA.

Shrivastav, R. & Sapienza, C.M. (2003). Objective measures of breathy voice quality obtained using an auditory model. *Journal of Acoustical Society of America, 114(4),* 2217-2224.

Singh, S. & Murry, T. (1978). Multidimensional classification of normal voice qualities. *Journal of Acoustical Society of America, 64(1),* 81-87.

Skodda, S. Visser, W. & Schlegel, Uwe. (2010). Gender-related patterns of dysprosody in Parkinson disease and correlation between speech variables and motor symptoms. *Journal of Voice, 25(1),* 76-82.

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

Sundberg, J. & Hogset, C. (2001). Voice source differences between falsetto and modal registers in counter tenors, tenors and baritones. *Log Phon Vocol (26),* 26-36.

Van Borsel, J. (2001). Physical appearance and voice in male-to-female transgenders. *Journal of the Voice, 15(4)*, 570-575.

Van Borsel, J., Eynde, E.V., Cuypere, G. & Bonte, K. (2008). Feminine after cricothyriod approximation? *Journal of Voice, 22(3)*, 379-384.

Van Borsel, J., Janssens, J. & De Bodt, M. (2009). Breathiness as a feminine voice characteristic: A perceptual approach. *Journal of the Voice, 23(3),* 291-294.

Vance, S. (2010). Opinions about the DSM gender identity disorder diagnosis: Result from an international administered to organizations concerned with the welfare of transgender people. *The International Journal of Transgenderism, 12(1),* 1-14.

Winters, K. (2011). The proposed gender dysphoria diagnosis in the DSM-5. *GID Reform Weblog by Kelly Winters.* http://gidreform.wordpress.com/2011/06/07/the-proposed-gender-dysphoria-diagnosis-in-the-dsm-5/.

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

Wu, K. & Childers, G. (1991). Gender recognition from speech. Part 1: Coarse analysis. *The Journal of the Acoustical Society of America, 90(4),* 1828-1840.

APPENDIX 1: RECORDED MATERIAL

1.1 Carrier Phrases:

1. Please say the word *heed* again
2. Please say the word *hid* again
3. Please say the word *hayed* again
4. Please say the word *head* again
5. Please say the word *had* again
6. Please say the word *heard* again
7. Please say the word *hud* again
8. Please say the word *hodd* again
9. Please say the word *hawed* again
10. Please say the word *hoed* again
11. Please say the word *hood* again
12. Please say the word *who'd* again
13. Please say the word *howd* again
14. Please say the word *hoyd* again
15. Please say the word *hide* again.

1.2 Excerpt from the Rainbow Passage – a public domain text widely used for language and speech studies:

When the sunlight strikes raindrops in the air, they act as a prism and form a rainbow. The rainbow is a division of white light into many beautiful colours. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow. Throughout the centuries people have explained the rainbow in various ways. Some have accepted it as a miracle without physical explanation. To the Hebrews it was a token that there would be no more universal floods. The Greeks used to imagine that it was a sign from the gods to foretell war or heavy rain. The Norsemen considered the rainbow as a bridge over which the gods passed from earth to their home in the sky. Others have tried to explain the phenomenon physically.

APPENDIX 2: RAW ACOUSTIC DATA

Gender	Mean f_0	H1-H2	H1-A1	H1-A3	F_1 BW
	Hz	dB	dB	dB	Hz
F1	200.00	6.78	0.23	24.91	169.78
F2	**261.47**	**16.47**	**26.41**	14.33	414.05
F3	177.19	2.84	-7.30	20.21	137.21
F4	171.70	0.10	-10.39	10.13	124.31
F5	209.11	4.27	23.68	23.68	172.06
F6	181.50	4.34	18.55	24.97	207.63
F7	169.82	2.20	22.79	18.71	131.47
F8	159.70	4.39	24.95	23.78	159.23
F9	156.98	11.09	-4.03	17.27	194.08
F10	239.41	5.60	14.85	16.00	93.28
F11	239.19	2.85	-10.77	*-3.39*	108.69
F12	169.08	0.85	12.02	26.70	151.61
F13	199.35	3.63	7.66	22.40	380.42
F14	205.03	4.92	20.26	28.92	281.84
F15	180.06	7.74	12.83	27.90	268.35
F16	204.01	11.28	22.95	34.65	259.30
F17	201.45	3.97	-3.09	24.69	131.31
F18	228.99	10.19	15.39	27.43	**601.64**
F19	212.54	6.17	-0.59	**36.50**	219.25
F20	158.62	3.78	-0.71	17.15	205.77
F21	227.14	0.57	-9.59	10.70	*72.90*
F22	227.15	4.80	5.99	20.51	328.68
F23	180.60	2.35	2.21	22.10	117.15
F24	172.89	6.42	-0.25	26.30	114.03
F25	157.79	7.60	7.83	33.51	381.87
F26	211.66	*-4.67*	-9.76	4.43	247.10
F27	*144.17*	2.57	-3.20	21.78	175.45
F28	192.75	1.47	-5.37	19.43	124.52
F29	174.87	4.30	2.34	35.57	272.15
F30	214.41	-4.15	*-13.91*	5.10	109.72
Mean	194.29	4.49	5.40	21.21	211.82
SD	± 29.15	± 4.30	± 12.41	± 9.32	± 117.05

Table 2.1a: Raw data with mean and standard deviation of the female participants for sustained vowel. Minimum values are italicized and maximum values are bolded

Gender	Jitter	Shimmer	NHR	VTI	SPI	CPP	CPPs
	%	%				dB	dB
F1	1.31	2.48	0.13	0.03	30.65	16.49	7.29
F2	0.35	*1.47*	0.12	0.04	7.27	22.20	7.43
F3	0.32	1.54	0.12	0.04	14.09	**24.17**	9.66
F4	0.50	2.70	0.13	**0.07**	9.98	23.24	**9.70**
F5	0.39	1.75	*0.11*	0.02	19.71	20.06	7.82
F6	0.95	3.98	0.14	0.03	17.28	17.77	6.95
F7	*0.29*	3.08	0.11	0.06	13.82	21.38	9.18
F8	0.63	2.00	0.11	0.05	26.75	18.81	8.45
F9	2.30	**8.66**	**0.28**	0.06	19.82	16.95	6.74
F10	0.36	2.07	0.11	0.05	17.12	21.60	7.73
F11	0.42	3.90	0.13	0.04	*6.08*	19.25	7.01
F12	0.65	3.64	0.14	0.06	10.52	21.18	8.36
F13	**4.49**	4.46	0.14	0.03	33.57	*11.97*	*4.47*
F14	0.55	3.25	0.11	0.04	12.82	18.01	7.02
F15	0.70	3.18	0.13	0.05	33.16	20.17	8.41
F16	0.88	3.86	0.13	0.03	39.67	15.09	6.36
F17	0.32	2.04	0.12	0.02	24.13	19.43	8.37
F18	4.21	6.34	0.15	0.05	17.99	12.17	4.54
F19	0.47	1.62	0.11	*0.01*	**43.11**	19.84	8.07
F20	0.68	3.19	0.15	0.05	14.71	21.06	8.59
F21	0.31	1.78	0.11	0.04	14.11	21.94	7.60
F22	3.57	6.93	0.19	0.06	6.67	16.28	5.36
F23	0.73	2.70	0.12	0.06	6.92	20.30	8.76
F24	0.40	2.88	0.12	0.03	41.30	19.38	8.89
F25	1.29	4.69	0.12	0.05	23.34	15.22	7.09
F26	1.17	4.11	0.14	0.04	19.57	20.24	7.36
F27	0.87	8.50	0.17	0.06	18.27	16.36	7.47
F28	0.47	2.99	0.12	0.05	14.28	20.80	8.38
F29	0.88	3.08	0.13	0.04	28.27	16.93	7.86
F30	1.46	4.22	0.16	0.03	36.87	15.60	6.05
Mean	1.06	3.57	0.13	0.04	20.73	18.80	7.57
SD	± 1.12	± 1.89	± 0.33	± 0.01	± 10.82	± 3.01	± 1.30

Table 2.1b: Raw data with mean and standard deviation of the female participants for sustained vowel. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	H1-H2	H1-A1	H1-A3	F_1 BW	Jitter	Shimmer	NHR	VTI	SPI
	Hz	dB	dB	dB	Hz	%	%			
F1	196.61	3.64	-10.23	15.18	**1043.89**	2.14	5.64	0.47	0.15	8.41
F2	209.93	6.94	1.28	13.72	880.07	1.06	5.77	0.46	0.10	2.59
F3	151.42	2.83	-6.12	*2.53*	200.46	1.62	3.85	0.62	0.08	2.69
F4	246.43	10.42	0.78	22.33	431.16	1.06	3.94	0.54	0.13	2.81
F5	148.31	3.85	-7.35	9.73	120.30	1.09	3.61	0.57	0.21	3.58
F6	206.41	6.55	-4.65	11.23	148.30	1.03	6.08	0.31	0.09	3.46
F7	191.00	4.10	-7.00	12.60	75.66	2.20	5.74	0.25	0.04	7.07
F8	165.38	0.31	-7.78	18.93	92.10	0.97	3.72	0.25	0.05	8.39
F9	151.30	0.31	*-11.89*	11.92	63.64	1.22	5.63	0.26	0.11	3.51
F10	168.49	0.82	-9.14	4.10	164.10	1.36	4.02	**0.71**	**0.29**	3.45
F11	**267.21**	6.91	-5.08	11.14	143.36	1.62	4.43	0.19	0.23	6.38
F12	210.22	0.89	-8.63	5.77	126.95	*0.53*	3.22	0.29	0.05	2.92
F13	*139.50*	4.17	-1.65	21.47	343.42	4.07	**8.54**	0.49	0.20	7.03
F14	241.12	8.39	-1.52	10.74	136.31	0.68	3.16	*0.13*	0.05	4.22
F15	205.81	9.06	**7.43**	25.47	268.01	1.65	4.17	0.16	0.11	6.25
F16	193.58	0.55	-5.94	15.19	118.32	3.38	5.75	0.43	0.16	4.44
F17	195.25	4.57	-1.49	13.87	148.28	0.88	4.48	0.21	*0.04*	6.90
F18	227.86	**12.58**	5.47	**26.54**	336.77	**6.82**	5.93	0.19	0.16	7.87
F19	217.92	7.74	1.92	22.65	122.40	0.85	*4.20*	0.15	0.05	5.61
F20	179.31	5.10	2.83	21.92	241.78	1.22	4.52	0.16	0.06	2.49
F21	177.58	*-2.24*	-10.50	8.69	65.20	1.04	2.83	0.13	0.06	6.32
F22	185.13	-0.33	-1.12	12.26	321.89	2.41	6.72	0.24	0.11	3.54
F23	177.83	3.57	-6.46	11.95	87.29	0.77	5.08	0.41	0.05	2.98
F24	174.03	1.06	-3.23	8.81	84.75	1.09	3.33	0.22	0.14	4.41
F25	174.21	4.12	5.12	22.45	262.89	2.88	3.84	0.17	0.06	6.73
F26	204.10	1.39	-9.06	12.89	137.00	1.99	5.10	0.25	0.11	3.84
F27	163.28	3.88	0.92	21.44	139.94	1.29	7.59	0.19	0.10	5.54
F28	213.30	4.98	2.29	24.62	201.42	1.24	4.00	0.17	0.06	4.92
F29	161.37	2.53	6.53	18.84	128.97	1.41	3.22	0.13	0.06	6.12
F30	181.05	2.38	-3.69	25.52	*51.61*	3.88	8.06	0.35	0.05	10.71
Mean	194.29	4.04	5.40	21.21	211.82	1.06	3.57	0.13	0.04	20.73
SD	± 29.15	± 3.42	± 12.41	± 9.32	±117.05	± 1.12	± 1.89	± 0.33	± 0.01	± 10.82

Table 2.2: Raw data with mean and standard deviation of the female participants for isolated /æ/ from *had*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	SD (f_0)	Range of Semitones	SD (semitones)	highest freq	lowest freq	CPP	CPPs
	Hz	Hz			Hz	Hz	Hz	Hz
F1	203.11	63.89	17	4.91	347.22	129.53	12.23	4.44
F2	196.57	34.42	12	2.9	284.09	138.12	12.15	3.93
F3	155.1	47.39	16	4.92	255.1	102.04	13.78	4.97
F4	211.6	66.04	19	**5.05**	352.11	118.48	13.04	4.61
F5	175.19	40.49	15	3.84	265.96	111.61	12.21	4
F6	194.73	33.33	14	2.84	274.73	126.9	12.86	4.68
F7	219.11	**66.44**	16	4.98	328.95	127.55	13.26	4.4
F8	168.78	23.21	10	2.24	225.23	126.26	12.17	3.95
F9	*149.23*	15.77	9	1.94	182.48	109.17	12.31	4.25
F10	207.09	58.5	16	4.57	328.95	128.21	12.81	4.07
F11	**236.64**	47.83	13	3.45	333.33	159.24	12.46	4.21
F12	196.53	40.84	15	3.49	328.95	138.12	14.27	**5.36**
F13	168.91	33.38	13	3.19	257.73	125	11.39	3.96
F14	213.28	52.45	14	4.01	337.84	149.7	12.7	4.24
F15	204.97	58.79	17	4.65	328.95	124.38	11.76	3.89
F16	172.46	30.85	18	3.07	255.1	91.24	12.72	4.56
F17	199.88	37.33	11	3.04	287.36	156.25	12.87	4.4
F18	217.03	30.65	8	2.46	274.73	170.07	*10.32*	*3*
F19	188.93	33.9	18	3.44	263.16	92.59	12.57	4.13
F20	183.65	35.22	14	3.02	312.5	142.05	14.15	5.11
F21	171.11	30.12	20	2.77	357.14	109.17	14.65	5.01
F22	217.47	45.81	11	3.39	312.5	166.67	13.29	4.31
F23	177.29	50.14	15	4.39	297.62	122.55	13.11	4.72
F24	171.98	15.34	*6*	1.5	213.68	143.68	13.8	5.01
F25	181.32	19.57	7	1.8	229.36	153.37	11.92	3.91
F26	200.86	48.11	**28**	3.95	384.62	77.88	13.83	4.58
F27	161.63	16.43	7	1.77	198.41	131.58	12.01	3.94
F28	213.66	33.81	15	2.59	316.46	133.69	12.61	4.36
F29	162.1	*12.66*	10	*1.32*	247.52	135.87	**14.73**	5.3
F30	191.45	26.89	17	2.7	260.42	99.21	12.05	3.96
Mean	190.39	38.32	14.03	3.27	288.07	128.01	12.8	4.38
SD	±21.98	±15.42	±4.59	±1.09	± 50.55	± 22.48	±0.98	±0.51

Table 2.3: Raw data with mean and standard deviation of the female participants for measurements from entire phrase with *had*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	H1-H2	H1-A1	H1-A3	F_1 BW	Jitter	Shimmer	NHR	VTI	SPI
	Hz	dB	dB	dB	Hz	%	%			
F1	222.36	4.66	-4.59	6.90	100.82	1.60	3.23	0.64	*0.01*	23.81
F2	**287.63**	8.03	-3.05	8.18	**309.18**	2.12	8.30	**0.83**	0.09	12.39
F3	158.95	3.87	-0.52	15.89	142.43	1.88	3.58	0.44	0.04	12.26
F4	241.82	**11.65**	-6.53	20.08	191.78	0.89	3.13	0.17	0.03	18.69
F5	*151.06*	3.13	*-12.74*	10.63	133.45	1.54	7.60	0.52	0.04	23.38
F6	200.40	4.76	0.51	15.09	169.19	0.71	3.40	0.26	0.06	15.99
F7	270.27	6.96	-8.90	16.77	113.48	1.25	6.64	0.57	0.07	18.35
F8	164.03	0.70	-4.81	17.94	92.57	0.87	*2.09*	0.25	0.05	26.13
F9	163.22	1.33	-9.37	9.91	*53.12*	1.04	6.13	0.22	0.07	19.10
F10	185.53	0.91	-8.09	*6.72*	144.31	1.85	6.42	0.56	**0.11**	17.17
F11	284.36	4.73	-8.03	16.94	127.03	**4.11**	**8.40**	0.52	0.03	23.17
F12	233.74	5.91	-11.01	8.83	206.60	1.01	3.18	0.57	0.04	25.80
F13	172.97	2.49	-0.13	32.29	215.21	1.77	5.76	0.19	0.05	29.32
F14	225.03	4.71	-4.98	14.62	172.16	0.58	3.50	0.14	0.04	14.50
F15	212.72	5.04	2.24	**33.49**	196.40	2.56	4.43	0.53	0.03	33.21
F16	202.74	2.04	-6.40	21.80	143.35	1.48	5.36	0.13	0.04	28.60
F17	178.53	2.65	-2.05	19.47	147.26	0.64	2.94	0.16	0.05	27.92
F18	217.06	5.56	**7.09**	32.35	234.61	0.91	3.45	*0.10*	0.04	40.83
F19	202.60	2.02	-9.66	15.28	88.86	2.68	7.82	0.41	0.04	28.44
F20	178.20	3.46	-4.22	14.67	158.92	0.52	2.63	0.15	0.05	14.50
F21	172.16	*-2.34*	-9.19	15.04	69.90	*0.39*	3.96	0.12	0.04	26.93
F22	202.06	2.77	-6.65	17.57	244.62	1.19	4.37	0.15	0.04	15.31
F23	182.24	0.35	-7.61	8.23	187.76	0.87	4.87	0.26	0.07	*8.68*
F24	178.99	1.72	-6.66	13.48	124.13	1.10	5.19	0.18	0.08	18.74
F25	174.64	4.24	4.06	29.40	230.83	0.64	3.15	0.16	0.05	32.12
F26	214.98	1.70	-6.00	13.41	110.07	1.10	5.76	0.14	0.05	17.66
F27	174.95	3.67	-2.02	18.69	143.59	0.87	5.37	0.13	0.06	13.73
F28	206.69	2.45	-3.48	31.25	214.50	0.74	3.72	0.14	0.04	27.51
F29	156.80	1.66	-0.92	25.04	291.02	0.55	3.59	0.13	0.06	11.67
F30	196.94	1.88	0.38	26.87	184.99	0.79	4.65	0.17	0.05	**43.91**
Mean	200.46	3.42	-4.44	17.89	164.74	1.27	4.75	0.30	0.05	22.33
SD	± 36.09	± 2.63	± 4.63	± 7.98	± 61.67	± 0.80	± 1.76	± 0.20	± 0.02	± 8.56

Table 2.4: Raw data with mean and standard deviation of the female participants for isolated /ɒ/ from *hodd*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	SD (f_0)	Range of Semitones	SD (semitones)	highest freq	lowest freq	CPP	CPPs
	Hz	Hz			Hz	Hz	dB	dB
F1	214.05	61.61	16.00	4.91	333.33	132.28	13.07	4.60
F2	205.68	54.98	14.00	4.19	333.33	145.35	12.17	4.16
F3	*146.54*	47.59	19.00	5.33	242.72	80.65	13.53	5.07
F4	200.73	61.71	18.00	5.04	324.68	118.48	13.13	5.02
F5	168.97	35.75	17.00	3.47	245.1	91.58	11.25	*3.13*
F6	192.30	32.81	20.00	3.30	250	78.62	12.49	4.23
F7	225.83	57.15	18.00	4.56	320.51	118.48	13.80	4.90
F8	168.61	26.22	12.00	2.53	245.1	126.26	12.72	4.27
F9	153.52	15.55	8.00	1.79	183.82	118.48	11.86	3.93
F10	216.16	49.69	14.00	3.76	324.68	147.93	12.80	4.02
F11	**240.66**	52.50	12.00	3.72	337.84	163.4	12.28	4.29
F12	193.50	41.91	14.00	3.64	316.46	140.45	**15.10**	**5.94**
F13	191.37	40.00	**25.00**	3.67	287.36	71.23	11.34	3.92
F14	211.50	44.84	13.00	3.54	333.33	159.24	12.66	4.22
F15	203.12	**74.33**	20.00	**5.70**	367.65	113.64	11.36	3.75
F16	176.98	23.76	8.00	2.25	217.39	138.12	12.53	4.36
F17	194.42	28.45	10.00	2.53	260.42	149.7	13.60	4.82
F18	217.20	22.88	9.00	1.99	274.73	164.47	*10.92*	3.57
F19	196.87	35.79	13.00	3.01	284.09	132.98	12.53	4.18
F20	185.36	49.80	15.00	4.04	308.64	132.28	13.55	4.91
F21	174.13	12.75	8.00	1.31	219.3	138.12	14.03	4.89
F22	200.73	37.88	23.00	3.64	277.78	72.89	12.60	4.09
F23	173.87	54.70	23.00	4.97	290.7	75.76	12.58	4.35
F24	173.86	16.47	8.00	1.56	227.27	147.93	12.75	4.42
F25	184.54	20.80	*7.00*	1.90	233.64	156.25	12.47	4.39
F26	203.90	26.40	19.00	2.44	263.16	86.21	14.26	4.87
F27	168.71	10.29	15.00	1.46	200	82.78	12.93	4.43
F28	206.56	22.59	12.00	1.85	297.62	145.35	13.37	4.87
F29	158.25	*9.28*	*7.00*	*0.98*	206.61	135.14	14.37	5.12
F30	198.68	17.39	*7.00*	1.50	263.16	170.07	12.07	4.06
Mean	191.55	36.20	14.13	3.15	275.68	124.47	12.80	4.43
SD	±22.05	±17.53	± 5.24	± 1.33	± 47.91	± 30.72	±0.97	±0.55

Table 2.5: Raw data with mean and standard deviation of the female participants for measurements from entire phrase with *hodd*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	H1-H2	H1-A1	H1-A3	F_1 BW
	Hz	dB	dB	dB	Hz
M31	133.05	-0.09	-10.37	14.13	106.85
M32	146.03	3.51	9.53	15.31	157.40
M33	164.36	3.22	-7.11	10.48	80.32
M34	128.66	3.01	-5.69	20.73	58.80
M35	120.81	2.55	-6.51	8.36	101.13
M36	110.58	1.60	-10.80	*4.00*	59.28
M37	120.25	4.26	-2.99	23.58	75.15
M38	*79.22*	0.32	-7.13	17.61	*56.81*
M39	185.77	**9.53**	1.71	29.16	213.62
M40	157.42	-1.38	-11.06	11.05	86.42
M41	135.88	2.50	3.38	29.81	105.74
M42	132.76	2.80	-5.25	11.74	120.32
M43	148.45	2.20	-3.47	15.66	83.41
M44	168.36	2.66	**13.59**	31.03	104.89
M45	104.91	3.00	-7.83	10.51	86.02
M46	121.21	4.96	-3.81	18.05	129.22
M47	108.16	*-3.30*	-1.24	19.51	214.92
M48	**186.94**	9.27	8.64	**40.00**	155.60
M49	133.06	0.69	-3.00	15.97	124.55
M50	104.23	-2.71	-4.90	11.12	111.62
M51	104.45	0.99	-2.96	26.62	99.59
M52	133.85	8.38	10.10	20.43	**298.73**
M53	113.76	-2.71	*-11.59*	13.74	64.28
M54	108.35	2.43	-4.24	13.94	154.37
M55	103.52	-2.24	-11.47	12.84	76.09
M56	96.52	0.60	-8.33	12.53	95.79
M57	120.70	0.73	1.46	23.49	258.48
M58	129.60	1.28	-5.75	11.48	107.37
Mean	128.60	2.07	-3.11	17.60	120.96
SD	± 26.29	± 3.26	± 6.87	± 8.09	± 60.45

Table 2.6a: Raw data with mean and standard deviation of the male participants for sustained vowel. Minimum values are italicized and maximum values are bolded

Gender	Jitter	Shimmer	NHR	VTI	SPI	CPP	CPPs
	%	%				dB	dB
M31	*0.26*	2.28	0.12	0.05	15.26	21.39	10.40
M32	0.31	2.09	0.12	0.03	8.05	24.00	10.66
M33	0.97	2.60	0.13	0.03	10.90	22.32	8.74
M34	0.70	3.72	0.13	0.03	31.37	18.31	8.76
M35	0.49	3.68	0.14	0.06	8.49	23.13	11.60
M36	1.17	5.86	0.18	0.05	11.08	18.15	9.24
M37	0.46	4.26	0.13	0.04	30.11	20.65	10.22
M38	1.24	6.86	0.17	0.04	19.86	16.65	8.40
M39	0.91	4.62	0.13	0.02	28.26	17.79	6.79
M40	1.06	1.94	0.12	0.05	11.01	25.00	10.79
M41	0.48	3.25	0.12	0.05	**40.61**	17.90	9.11
M42	0.44	*1.86*	*0.11*	0.05	17.77	**26.15**	11.89
M43	0.57	2.00	0.13	0.03	12.98	22.88	12.04
M44	0.94	4.06	0.14	0.05	18.63	16.63	9.68
M45	0.71	6.20	0.16	0.05	17.63	19.30	10.55
M46	0.27	3.29	0.12	*0.01*	16.63	24.44	**12.93**
M47	0.42	2.27	0.14	0.07	15.74	19.08	9.97
M48	0.71	2.56	0.13	0.03	36.94	16.66	*6.48*
M49	0.52	2.72	0.14	0.02	10.05	22.01	12.21
M50	**1.52**	**7.94**	0.18	0.05	9.97	17.80	9.00
M51	0.39	3.08	0.14	0.05	12.02	23.25	11.60
M52	1.25	7.50	0.13	0.05	13.66	*16.15*	7.58
M53	0.33	4.20	0.14	0.04	13.17	22.89	10.77
M54	0.71	7.04	0.17	**0.08**	*5.46*	18.35	8.35
M55	0.92	5.34	**0.20**	0.06	8.60	18.33	8.87
M56	0.69	4.58	0.16	0.05	15.37	21.96	11.12
M57	0.55	3.17	0.13	0.06	17.09	22.07	10.22
M58	0.82	2.54	0.16	0.04	19.33	21.39	9.62
Mean	0.71	3.98	0.14	0.04	17.00	20.52	9.91
SD	± 0.33	± 1.82	± 0.02	± .02	± 8.85	± 2.89	± 1.62

Table 2.6b: Raw data with mean and standard deviation of the male participants for sustained vowel. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	H1-H2	H1-A1	H1-A3	F_1 BW	Jitter	Shimmer	NHR	VTI	SPI
	Hz	dB	dB	dB	Hz	%	%			
M31	116.56	3.05	-7.63	5.88	43.25	1.52	7.17	0.38	0.07	3.26
M32	172.96	1.21	4.81	20.05	105.51	0.70	3.15	0.24	0.15	4.78
M33	192.58	4.18	-12.67	*-3.92*	79.49	1.20	3.06	0.52	0.07	2.48
M34	118.81	1.41	-3.27	19.38	*41.61*	1.75	7.32	*0.13*	0.13	5.44
M35	118.91	4.93	-12.34	2.81	52.86	3.47	7.85	**0.76**	0.17	2.64
M36	142.94	6.48	0.16	8.66	74.77	2.55	13.54	0.45	0.24	3.02
M37	120.93	3.95	-3.36	21.70	72.20	1.21	6.96	0.26	0.08	**14.18**
M38	106.10	1.94	-1.74	18.24	55.74	3.79	8.39	0.49	0.13	6.86
M39	**239.59**	14.83	5.80	**32.25**	236.11	4.31	7.12	0.29	0.10	10.25
M40	171.23	5.54	-1.89	13.29	**282.44**	1.56	2.69	0.48	0.09	3.18
M41	175.53	3.01	1.16	18.18	159.83	**6.60**	8.42	0.50	0.15	4.46
M42	124.72	3.03	-2.23	4.84	207.04	5.03	10.83	0.57	0.30	2.34
M43	98.27	*-7.14*	*-19.00*	-2.96	51.13	1.73	5.95	0.56	0.11	2.40
M44	164.17	4.64	-6.60	15.93	96.64	1.42	6.56	0.67	0.21	4.02
M45	104.49	3.07	-4.04	22.40	90.21	1.70	6.30	0.20	0.07	8.48
M46	126.40	0.37	-7.08	3.92	129.43	2.48	**15.53**	0.72	**0.33**	2.35
M47	114.17	-3.80	-1.35	15.82	91.89	1.27	6.04	0.22	0.10	5.73
M48	130.27	7.34	-1.71	17.83	83.15	2.08	14.66	0.32	0.12	5.69
M49	136.42	-0.64	-7.90	6.09	99.61	0.99	4.15	0.23	0.06	3.81
M50	*98.16*	-3.78	-10.37	5.72	84.14	*0.48*	8.75	0.18	0.13	*2.09*
M51	109.23	1.41	-5.20	14.33	98.52	1.59	1.51	0.21	0.07	4.22
M52	144.28	**16.57**	1.93	19.44	112.23	1.56	*0.71*	0.19	0.09	5.21
M53	110.79	-2.77	-13.03	4.81	58.18	6.09	5.58	0.68	0.26	2.27
M54	113.31	3.87	**6.28**	14.13	183.68	1.56	9.65	0.20	0.08	4.03
M55	101.07	-1.14	-12.47	4.15	64.40	0.85	6.02	0.18	0.08	4.57
M56	99.59	-4.57	-13.32	11.82	54.10	2.48	10.03	0.43	0.14	3.71
M57	127.26	1.29	-6.76	15.54	93.86	3.87	6.94	0.27	*0.04*	5.92
M58	127.26	7.28	-1.48	7.62	99.46	4.35	14.43	0.53	0.28	2.85
Mean	132.34	2.70	-4.83	12.07	103.62	2.43	7.47	0.39	0.14	4.65
SD	± 33.26	± 5.17	± 6.31	± 8.36	± 59.45	± 1.63	± 3.82	± 0.19	± 0.08	± 2.69

Table 2.7: Raw data with mean and standard deviation of the male participants for isolated /æ/ from *had*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	SD (f_0)	Range of Semitones	SD (semitones)	highest freq	lowest freq	CPP	CPPs
	Hz	Hz			Hz	Hz	dB	dB
M31	131.93	**45.63**	**26.00**	4.76	**378.79**	81.17	12.04	4.09
M32	167.59	44.24	14.00	4.42	247.52	107.30	11.35	3.49
M33	170.99	24.57	6.00	2.30	213.68	**142.86**	12.23	4.36
M34	109.25	9.44	7.00	1.52	128.87	88.03	11.67	4.35
M35	118.63	36.95	17.00	5.03	193.80	72.89	13.67	**5.39**
M36	128.60	18.53	10.00	2.50	171.23	100.81	11.54	4.04
M37	126.81	12.26	8.00	1.72	164.47	104.60	11.32	3.69
M38	*94.36*	17.72	11.00	3.15	130.21	*70.03*	10.95	3.50
M39	**182.30**	34.51	12.00	3.14	257.73	131.58	11.35	4.18
M40	151.56	39.14	19.00	4.66	227.27	79.87	12.64	4.17
M41	144.60	23.31	12.00	2.67	204.92	106.84	*10.64*	*3.37*
M42	117.83	21.60	13.00	3.06	193.80	91.24	12.05	4.14
M43	104.91	35.11	16.00	**5.04**	189.39	73.10	12.68	4.52
M44	145.61	29.15	21.00	3.33	284.09	82.51	11.44	3.59
M45	115.55	21.11	10.00	3.05	162.34	89.93	11.67	4.41
M46	132.67	19.32	10.00	2.49	187.97	102.04	12.69	4.52
M47	105.64	10.62	9.00	1.77	128.87	79.37	11.28	3.57
M48	145.48	26.06	11.00	3.00	196.85	106.84	10.94	3.77
M49	133.76	13.80	10.00	1.89	179.86	105.49	12.71	4.37
M50	118.52	38.92	15.00	4.68	211.86	87.11	12.94	4.69
M51	108.06	5.61	6.00	*0.91*	138.12	96.53	13.22	4.95
M52	134.96	18.27	15.00	2.16	247.52	102.88	11.24	3.88
M53	125.97	16.18	13.00	2.18	185.19	88.03	12.80	4.86
M54	120.31	19.89	10.00	2.90	155.28	85.62	12.65	4.99
M55	99.70	*5.25*	*3.00*	1.00	*113.12*	91.91	**14.24**	4.93
M56	101.46	40.30	24.00	3.86	347.22	86.21	11.46	3.87
M57	131.41	9.14	7.00	1.22	167.79	108.70	13.00	4.25
M58	139.59	28.23	15.00	3.66	190.84	83.06	11.45	3.90
Mean	128.86	23.75	12.50	2.93	199.95	94.52	12.07	4.21
SD	±21.92	±11.87	± 5.40	± 1.23	± 61.95	± 16.69	±0.90	±0.52

Table 2.8: Raw data with mean and standard deviation of the male participants for measurements from entire phrase with *had*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	H1-H2	H1-A1	H1-A3	F_1 BW	Jitter	Shimmer	NHR	VTI	SPI
	Hz	dB	dB	dB	Hz	%	%			
M31	142.04	5.02	-10.13	14.50	55.56	1.48	6.11	0.58	0.05	22.61
M32	182.68	2.83	-8.71	15.69	92.01	1.43	4.12	0.21	0.07	16.48
M33	**188.96**	2.40	-13.75	2.76	94.93	0.98	*0.46*	0.37	0.06	11.02
M34	119.63	1.27	-7.29	20.23	69.12	1.65	6.14	0.33	0.05	28.42
M35	120.34	2.95	-13.50	4.51	125.38	3.25	8.68	**0.90**	0.22	*4.75*
M36	168.29	3.82	-9.73	7.25	69.95	1.80	7.50	0.14	0.11	15.26
M37	127.35	3.31	-6.39	24.24	*47.07*	1.40	5.42	0.26	0.05	29.38
M38	*95.89*	*-5.44*	-13.44	26.56	74.44	12.09	**26.99**	0.82	0.05	23.08
M39	173.24	11.86	**1.80**	**35.90**	143.64	0.98	4.91	0.37	0.06	11.02
M40	159.05	2.97	-6.88	14.22	142.16	2.12	3.89	0.18	0.03	21.14
M41	158.16	1.85	-3.84	29.92	177.86	**13.63**	11.35	0.37	*0.02*	**33.46**
M42	125.21	1.69	-8.04	10.69	132.24	1.46	3.73	0.42	0.06	9.01
M43	116.32	-0.42	-13.29	8.47	73.72	1.70	5.89	0.53	0.05	12.74
M44	134.18	5.22	-8.10	9.11	112.77	1.58	4.17	0.31	0.11	11.39
M45	108.91	3.67	-6.77	13.44	74.24	1.11	7.85	0.23	0.05	18.33
M46	135.90	2.70	-8.58	8.84	235.88	1.57	8.63	0.54	0.06	14.17
M47	118.12	-3.82	-3.71	21.41	123.56	1.27	5.66	*0.12*	0.05	24.74
M48	137.65	7.46	0.97	25.87	82.74	0.84	7.64	0.16	0.04	22.53
M49	144.30	0.55	-6.94	14.18	91.55	3.85	10.05	0.39	0.04	20.87
M50	102.87	-4.04	-9.99	8.44	157.65	0.84	11.97	0.16	0.05	11.91
M51	106.51	1.82	-6.62	7.87	116.77	2.35	17.14	0.32	**0.24**	5.10
M52	146.81	**12.45**	0.23	28.60	115.23	2.56	7.01	0.19	0.02	19.63
M53	110.07	-1.10	-12.64	13.03	65.52	1.36	8.86	0.53	0.08	14.88
M54	113.97	2.97	-0.31	16.13	**414.66**	2.07	8.76	0.16	0.04	13.91
M55	99.11	-5.18	*-18.37*	*-3.43*	159.54	1.38	6.74	0.18	0.05	8.30
M56	97.97	-5.37	-16.95	8.44	89.25	4.51	6.17	0.38	0.06	22.09
M57	125.73	-0.42	-7.69	17.26	68.69	*0.63*	4.35	0.13	0.04	17.80
M58	125.73	6.17	-11.24	7.85	134.66	8.34	14.00	0.54	0.07	9.74
Mean	131.61	2.04	-8.21	14.71	119.31	2.79	8.01	0.35	0.07	16.92
SD	± 25.83	± 4.47	± 5.09	± 9.12	± 71.79	± 3.23	± 5.05	± 0.20	± 0.05	± 7.30

Table 2.9: Raw data with mean and standard deviation of the male participants for isolated /ɒ/ from *hodd*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	SD (f_0)	Range of Semitones	SD (semitones)	highest freq	lowest freq	CPP	CPPs
	Hz	Hz			Hz	Hz	dB	dB
M31	133.60	46.85	24.00	4.62	**320.51**	80.65	12.03	4.08
M32	172.50	32.75	16.00	3.17	274.73	111.11	11.46	3.27
M33	174.74	27.13	8.00	2.65	215.52	135.87	12.46	4.44
M34	113.64	9.33	6.00	1.48	132.28	89.93	12.21	4.65
M35	119.45	42.06	17.00	5.61	203.25	79.11	13.25	5.06
M36	128.83	28.94	13.00	3.76	186.57	89.29	11.14	3.81
M37	127.31	10.66	6.00	1.54	150.60	105.04	10.92	3.60
M38	107.90	16.38	11.00	2.81	134.41	*70.42*	11.24	3.75
M39	**182.66**	21.48	10.00	2.02	245.10	**136.61**	11.24	3.88
M40	115.49	45.23	**25.00**	4.67	308.64	74.40	11.25	3.94
M41	146.46	24.39	13.00	2.66	233.64	111.61	*10.46*	*3.12*
M42	120.75	16.40	9.00	2.35	162.34	97.28	12.09	4.07
M43	*99.25*	16.52	10.00	2.76	129.53	73.31	12.72	4.66
M44	135.27	20.14	12.00	2.51	200.00	95.79	11.51	3.52
M45	118.76	25.09	14.00	3.38	179.86	81.17	11.83	4.32
M46	135.21	19.30	9.00	2.48	173.61	105.49	12.28	4.32
M47	109.98	10.65	8.00	1.78	127.55	84.18	11.99	4.23
M48	133.04	9.94	6.00	1.28	163.40	115.21	10.55	3.73
M49	138.14	17.14	10.00	2.24	171.23	98.43	12.70	4.68
M50	129.08	**53.64**	16.00	**6.12**	235.85	91.91	12.08	4.47
M51	105.57	6.38	5.00	1.03	127.55	96.90	**13.82**	**5.32**
M52	131.83	13.43	8.00	1.80	159.24	98.43	11.24	3.89
M53	124.52	25.10	16.00	3.39	204.92	82.24	13.58	5.07
M54	125.96	36.12	17.00	4.12	235.85	87.41	12.40	4.97
M55	100.73	*4.77*	*3.00*	*0.79*	*112.61*	93.63	13.37	4.51
M56	106.69	11.21	8.00	1.83	130.89	80.65	11.60	3.93
M57	124.29	10.14	8.00	1.42	163.40	102.46	12.60	4.27
M58	138.94	24.64	13.00	3.28	186.57	87.11	12.12	4.73
Mean	128.59	22.35	11.46	2.77	188.20	94.84	12.01	4.22
SD	±20.86	±12.97	± 5.30	± 1.34	± 54.87	± 16.63	±0.87	±0.55

Table 2.10: Raw data with mean and standard deviation of the male participants for measurements from entire phrase with *hodd*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	H1-H2	H1-A1	H1-A3	F_1 BW
	Hz	dB	dB	dB	Hz
T59	**248.73**	3.68	**2.72**	**28.57**	145.05
T60	187.53	**5.73**	-0.97	*12.08*	160.00
T61	155.10	3.20	-2.12	24.92	*104.38*
T62	*154.21*	5.24	*-2.32*	17.09	158.08
T63	161.89	*1.38*	0.24	25.10	**647.65**
Mean	181.49	3.84	-0.49	21.56	243.03
SD	± 39.95	± 1.73	± 2.07	± 6.76	± 227.30

Table 2.11a: Raw data with mean and standard deviation of the transgender/puberphonic voice group participants for sustained vowel. Minimum values are italicized and maximum values are bolded

Gender	Jitter	Shimmer	NHR	VTI	SPI	CPP	CPPs
	%	%				dB	dB
T59	0.42	2.94	*0.11*	0.04	16.71	19.19	6.08
T60	*0.22*	1.34	0.12	0.04	**40.70**	21.00	**12.38**
T61	1.11	5.85	**0.16**	0.04	21.27	16.55	*5.80*
T62	0.64	*0.23*	0.14	0.04	*12.87*	**23.03**	11.53
T63	**1.48**	**5.90**	0.14	**0.05**	20.35	*14.93*	7.78
Mean	0.77	3.25	0.14	0.04	22.38	18.94	8.71
SD	± 0.52	± 2.59	± 0.02	± 0.01	± 10.77	± 3.27	± 3.07

Table 2.11b: Raw data with mean and standard deviation of the transgender/puberphonic voice group participants for sustained vowel. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	H1-H2	H1-A1	H1-A3	F_1 BW	Jitter	Shimmer	NHR	VTI	SPI
	Hz	dB	dB	dB	Hz	%	%			
T59	**266.14**	7.40	**1.64**	**27.49**	**195.68**	*1.38*	4.73	**0.46**	**0.07**	*3.78*
T60	246.03	**13.19**	-2.88	*10.17*	131.91	2.97	*3.60*	0.19	0.06	5.17
T61	169.52	5.94	-2.94	24.09	161.39	**4.24**	**8.91**	0.26	0.06	6.54
T62	200.74	9.30	*-3.48*	15.92	194.11	2.06	5.88	0.34	*0.05*	5.98
T63	*156.30*	*3.92*	-0.53	24.33	*126.80*	1.48	6.49	*0.13*	0.06	**7.62**
Mean	207.74	7.95	-1.63	20.40	161.98	2.43	5.92	0.28	0.06	5.82
SD	± 47.52	± 3.53	± 2.16	± 7.14	± 32.83	± 1.19	± 2.01	± 0.13	± 0.01	± 1.45

Table 2.12: Raw data with mean and standard deviation of the transgender/puberphonic voice group participants for isolated /æ/ from *had*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	SD (f_0)	Range of Semitones	SD (semitones)	highest freq	lowest freq	CPP	CPPs
	Hz	Hz			Hz	Hz	dB	dB
T59	**239.47**	32.39	13.00	2.31	**342.47**	**168.92**	11.99	*2.49*
T60	219.52	37.55	**16.00**	3.39	274.73	112.61	12.26	3.92
T61	*144.33*	22.28	10.00	2.54	*192.31*	113.12	*11.40*	3.77
T62	182.64	**40.02**	**16.00**	**3.90**	255.10	*104.17*	**12.51**	**4.50**
T63	154.36	*10.20*	*6.00*	*1.12*	201.61	139.66	11.76	4.06
Mean	188.06	28.49	12.20	2.65	253.24	127.70	11.98	3.75
SD	±40.96	±12.28	±4.27	±1.07	±60.84	±26.63	±0.43	±0.75

Table 2.13: Raw data with mean and standard deviation of the transgender/puberphonic voice group participants for measurements from entire phrase with *had*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	H1-H2	H1-A1	H1-A3	F_1 BW	Jitter	Shimmer	NHR	VTI	SPI
	Hz	dB	dB	dB	Hz	%	%			
T59	**293.77**	7.93	0.58	**38.83**	232.40	1.02	*2.51*	*0.15*	0.04	23.67
T60	211.42	**8.28**	**4.22**	18.33	**776.78**	1.16	2.57	**0.22**	**0.07**	*12.43*
T61	174.47	5.93	-0.85	*16.24*	*66.41*	**1.20**	**6.43**	0.15	*0.03*	25.30
T62	167.59	2.58	*-4.98*	18.86	149.34	0.93	5.78	0.17	0.06	**26.34**
T63	*147.25*	*1.57*	-2.42	18.72	75.13	*0.92*	6.28	0.18	0.06	16.88
Mean	198.90	5.26	-0.69	22.19	260.01	1.05	4.71	0.17	0.05	20.93
SD	± 57.88	± 3.06	± 3.43	± 9.36	± 296.53	± 0.13	± 2.00	± 0.03	± 0.02	± 6.01

Table 2.14: Raw data with mean and standard deviation of the transgender/puberphonic voice group participants for isolated /ɒ/ from *hodd*. Minimum values are italicized and maximum values are bolded

Gender	Mean f_0	SD (f_0)	Range of Semitones	SD (semitones)	highest freq	lowest freq	CPP	CPPs
	Hz	Hz			Hz	Hz	dB	dB
T59	**232.79**	**40.17**	10.00	2.78	**312.50**	**172.41**	*11.29*	4.04
T60	208.66	31.26	12.00	2.83	257.73	132.28	11.80	*3.32*
T61	153.49	21.59	11.00	2.39	211.86	*113.12*	12.17	4.12
T62	157.89	38.33	**15.00**	**3.98**	255.10	*113.12*	**12.62**	**4.59**
T63	*149.68*	*7.28*	*7.00*	*0.96*	*181.16*	124.38	11.54	3.56
Mean	180.50	27.73	11.00	2.59	243.67	131.06	11.88	3.93
SD	±37.81	±13.56	±2.92	±1.09	±49.96	±24.49	±0.52	±0.50

Table 2.15: Raw data with mean and standard deviation of the transgender/puberphonic voice group participants for measurements from entire phrase with *hodd*. Minimum values are italicized and maximum values are bolded

APPENDIX 3: RESPONSE DATA

Table 3.1: Listener Response for carrier phrase with *had*
Note: Values one standard deviation below the mean are bolded and one standard deviation above are italicized

	Female	Male	Proportion	Gender rating
F1	26	0	100	**-85.92**
F2	25	1	96.15	*-47.23*
F3	26	0	100	-62.15
F4	26	0	100	**-85.31**
F5	26	0	100	-65.46
F6	26	0	100	-76.62
F7	26	0	100	**-90.58**
F8	26	0	100	-54.50
F9	25	1	96.15	*-25.31*
F10	26	0	100	-71.85
F11	26	0	100	-79.62
F12	26	0	100	-73.62
F13	26	0	100	-72.81
F14	26	0	100	-79.96
F15	25	0	96.15	-74.73
F16	26	0	100	*-38.58*
F17	25	1	96.15	*-44.62*
F18	26	0	100	-79.15
F19	26	0	100	-64.23
F20	26	0	100	-69.54
F21	26	0	100	-56.62
F22	26	0	100	**-87.85**
F23	26	0	100	-51.35

	Female	Male	Proportion	Gender rating
F24	26	0	100	-57.27
F25	26	0	100	-75.96
F26	26	0	100	-70.77
F27	26	0	100	-59.12
F28	26	0	100	-70.81
F29	22	4	**84.62**	*-27.69*
F30	26	0	100	-74.00
Mean			98.97	-65.77
SD			3.02	16.83
M31	0	26	100.00	75.62
M32	0	26	100.00	50.27
M33	**17**	**9**	**34.62**	**-24.27**
M34	0	26	100.00	86.96
M35	0	26	100.00	87.96
M36	0	26	100.00	78.77
M37	0	26	100.00	75.88
M38	0	26	100.00	*93.50*
M39	0	26	100.00	57.96
M40	1	25	96.15	47.35
M41	0	26	100.00	66.73
M42	0	26	100.00	54.92
M43	0	26	100.00	85.38
M44	3	23	88.46	**17.27**
M45	0	26	100.00	81.19
M46	0	26	100.00	49.12
M47	0	26	100.00	74.69
M48	0	26	100.00	76.38

	Female	Male	Proportion	Gender rating
M49	1	25	96.15	46.42
M50	0	26	100.00	62.85
M51	0	26	100.00	69.50
M52	0	26	100.00	79.46
M53	0	26	100.00	75.50
M54	0	26	100.00	*89.65*
M55	0	26	100.00	80.65
M56	0	26	100.00	82.85
M57	0	26	100.00	51.15
M58	0	26	100.00	57.69
Mean			96.98	65.41
SD			12.44	24.67
T59	15	11	*57.69*	**-21.42**
T60	0	26	**0.00**	11.12
T61	0	26	**0.00**	*50.38*
T62	15	11	*57.69*	-16.81
T63	7	19	26.92	10.04
Mean			28.46	6.66
SD			28.86	28.65

Table 3.2: Listener Response for carrier phrase with *hodd*
Note: Values one standard deviation below the mean are bolded and one standard deviation above are italicized.

	Female	Male	Proportion	Gender rating
F1	25	1	96.15	-77.85
F2	26	0	100.00	-58.31
F3	25	1	96.15	*-44.62*
F4	26	0	100.00	-80.00
F5	25	1	96.15	*-44.88*
F6	26	0	100.00	-69.04
F7	26	0	100.00	**-87.65**
F8	26	0	100.00	-53.62
F9	26	0	100.00	*-30.35*
F10	26	0	100.00	-75.77
F11	26	0	100.00	-75.46
F12	25	1	96.15	-78.00
F13	26	0	100.00	-79.00
F14	26	0	100.00	-77.04
F15	26	0	100.00	-80.62
F16	25	1	96.15	-48.12
F17	26	0	100.00	-48.77
F18	26	0	100.00	-74.92
F19	26	0	100.00	-71.62
F20	26	0	100.00	-74.04
F21	26	0	100.00	-49.69
F22	26	0	100.00	**-90.96**
F23	26	0	100.00	*-41.04*
F24	26	0	100.00	-64.08
F25	25	1	96.15	-60.73

	Female	Male	Proportion	Gender rating
F26	26	0	100.00	-71.19
F27	26	0	100.00	-47.85
F28	26	0	100.00	-80.85
F29	22	4	**84.62**	*-25.08*
F30	26	0	100.00	-60.27
Mean			98.72	-64.05
SD			3.09	17.16
M31	0	26	100.00	84.85
M32	1	25	96.15	30.96
M33	**18**	**8**	30.77	-29.12
M34	0	26	100.00	85.69
M35	0	26	100.00	84.31
M36	0	26	100.00	81.81
M37	0	26	100.00	76.96
M38	1	25	96.15	89.96
M39	0	26	100.00	70.62
M40	0	26	100.00	76.92
M41	0	26	100.00	67.50
M42	0	26	100.00	55.00
M43	0	26	100.00	86.04
M44	0	26	100.00	35.19
M45	0	26	100.00	79.19
M46	0	26	100.00	45.12
M47	0	26	100.00	63.62
M48	0	26	100.00	74.81
M49	0	26	100.00	49.69
M50	0	26	100.00	54.65

	Female	Male	Proportion	Gender rating
M51	0	26	100.00	63.96
M52	0	26	100.00	86.54
M53	0	26	100.00	82.38
M54	0	26	100.00	78.65
M55	0	26	100.00	83.73
M56	0	26	100.00	81.00
M57	0	26	100.00	55.77
M58	0	26	100.00	51.62
Mean			97.25	65.98
SD			13.07	24.87
T59	9	17	*34.62*	**-10.35**
T60	2	24	7.69	11.15
T61	0	26	**0.00**	*36.77*
T62	2	24	7.69	31.23
T63	10	16	*38.46*	1.54
Mean			17.69	14.07
SD			17.54	19.82

Table 3.3: Listener Response for sustained vowel
Note: Values one standard deviation below the mean are bolded and one standard deviation above are italicized.

	Female	Male	Proportion	Gender rating	Overall rating
F1	23	3	88.46	-36.50	-46.56
F1	24	2	92.31	-48.12	
F1	26	0	*100.00*	**-55.08**	
F2	21	5	80.77	-29.54	-41.51
F2	22	4	84.62	-46.35	
F2	23	3	88.46	-48.65	
F3	20	6	76.92	-27.92	-24.40
F3	20	6	76.92	-31.31	
F3	17	9	65.38	-13.96	
F4	20	6	76.92	-25.73	-14.40
F4	16	10	61.54	-9.19	
F4	16	10	61.54	-8.27	
F5	20	6	76.92	-21.27	-21.87
F5	20	6	76.92	-31.31	
F5	16	10	61.54	-13.04	
F6	19	7	73.08	-27.15	-21.41
F6	17	9	65.38	-15.58	
F6	21	5	80.77	-21.50	
F7	19	7	73.08	-12.50	-15.54
F7	19	7	73.08	-23.92	
F7	17	9	65.38	-10.19	
F8	4	22	**15.38**	*27.85*	**26.19**
F8	6	20	**23.08**	*20.27*	
F8	7	19	**26.92**	*30.46*	
F9	18	8	69.23	-2.27	-11.40

	Female	Male	Proportion	Gender rating	Overall rating
F9	24	2	92.31	-26.35	
F9	21	5	80.77	-5.58	
F10	2	24	**7.69**	*17.77*	
F10	2	24	**7.69**	*24.38*	
F10	4	22	**15.38**	*11.69*	**17.95**
F11	26	0	*100.00*	**-60.31**	
F11	26	0	*100.00*	**-70.69**	
F11	25	1	96.15	**-60.27**	*-63.76*
F12	21	5	80.77	-28.58	
F12	21	5	80.77	-34.46	
F12	22	4	84.62	-30.35	-31.13
F13	23	3	88.46	**-52.58**	
F13	24	2	92.31	-43.31	
F13	24	2	92.31	**-52.65**	-49.51
F14	24	2	92.31	**-59.65**	
F14	23	3	88.46	**-56.19**	
F14	24	2	92.31	**-53.73**	*-56.53*
F15	22	4	84.62	-32.58	
F15	22	4	84.62	-39.92	
F15	22	4	84.62	**-52.92**	-41.81
F16	20	6	76.92	-28.27	
F16	19	7	73.08	-33.77	
F16	18	8	69.23	-19.69	-27.24
F17	17	9	65.38	-14.15	
F17	13	13	50.00	-4.65	
F17	10	16	**38.46**	-3.69	-7.50
F18	24	2	92.31	**-53.65**	*-54.05*

	Female	Male	Proportion	Gender rating	Overall rating
F18	23	3	88.46	**-51.19**	
F18	23	3	88.46	**-57.31**	
F19	23	3	88.46	-47.19	-40.56
F19	24	2	92.31	-43.96	
F19	21	5	80.77	-30.54	
F20	20	6	76.92	-27.58	-23.44
F20	21	5	80.77	-23.65	
F20	19	7	73.08	-19.08	
F21	7	19	**26.92**	*9.77*	**14.05**
F21	3	23	**11.54**	*15.04*	
F21	5	21	**19.23**	*17.35*	
F22	24	2	92.31	**-59.77**	*-59.99*
F22	25	1	96.15	**-58.27**	
F22	25	1	96.15	**-61.92**	
F23	23	3	88.46	-31.42	-33.56
F23	22	4	84.62	-34.81	
F23	23	3	88.46	-34.46	
F24	1	25	**3.85**	*32.19*	**34.17**
F24	1	25	**3.85**	*36.00*	
F24	0	26	**0.00**	*34.31*	
F25	20	6	76.92	-20.46	-17.81
F25	18	8	69.23	-23.85	
F25	17	9	65.38	-9.12	
F26	18	8	69.23	-22.85	-29.08
F26	18	8	69.23	-28.81	
F26	20	6	76.92	-35.58	
F27	15	11	57.69	-3.92	**9.47**

	Female	Male	Proportion	Gender rating	Overall rating
F27	11	15	42.31	*17.96*	
F27	10	16	**38.46**	*14.38*	
F28	22	4	84.62	-39.38	-47.10
F28	24	2	92.31	-49.54	
F28	24	2	92.31	**-52.38**	
F29	6	20	**23.08**	*15.58*	**19.56**
F29	4	22	**15.38**	*25.38*	
F29	6	20	**23.08**	*17.73*	
F30	23	3	88.46	-46.65	-50.01
F30	25	1	96.15	**-59.62**	
F30	25	1	96.15	-43.77	
Mean			68.93	-23.63	
SD			27.64	27.30	
M31	0	26	100.00	65.88	65.19
M31	0	26	100.00	59.85	
M31	0	26	100.00	69.85	
M32	0	26	100.00	60.77	65.79
M32	0	26	100.00	69.42	
M32	0	26	100.00	67.19	
M33	4	22	**84.62**	**20.42**	**25.14**
M33	1	25	96.15	**29.00**	
M33	2	24	92.31	**26.00**	
M34	0	26	100.00	70.65	73.69
M34	0	26	100.00	68.81	
M34	0	26	100.00	81.62	
M35	0	26	100.00	74.35	72.74
M35	0	26	100.00	73.92	

	Female	Male	Proportion	Gender rating	Overall rating
M35	0	26	100.00	69.96	
M36	0	26	100.00	84.19	79.15
M36	0	26	100.00	77.12	
M36	1	25	96.15	76.15	
M37	1	25	96.15	75.58	78.12
M37	1	25	96.15	76.12	
M37	0	26	100.00	82.65	
M38	0	26	100.00	*89.31*	*91.94*
M38	0	26	100.00	*92.35*	
M38	0	26	100.00	*94.15*	
M39	1	25	96.15	**37.15**	**25.19**
M39	3	23	88.46	**20.42**	
M39	4	22	**84.62**	**18.00**	
M40	1	25	96.15	54.88	50.44
M40	1	25	96.15	56.00	
M40	3	23	88.46	**40.42**	
M41	0	26	100.00	58.12	63.49
M41	0	26	100.00	61.58	
M41	0	26	100.00	70.77	
M42	0	26	100.00	66.27	65.95
M42	0	26	100.00	68.15	
M42	0	26	100.00	63.42	
M43	0	26	100.00	49.81	55.12
M43	0	26	100.00	60.58	
M43	0	26	100.00	54.96	
M44	12	14	**53.85**	**0.04**	**2.35**
M44	12	14	**53.85**	**0.62**	

	Female	Male	Proportion	Gender rating	Overall rating
M44	7	19	**73.08**	**6.38**	
M45	1	25	96.15	65.96	77.17
M45	0	26	100.00	83.85	
M45	0	26	100.00	81.69	
M46	0	26	100.00	72.46	72.12
M46	0	26	100.00	68.96	
M46	0	26	100.00	74.92	
M47	0	26	100.00	74.65	73.18
M47	0	26	100.00	69.38	
M47	0	26	100.00	75.50	
M48	4	22	**84.62**	**20.08**	**24.44**
M48	3	23	88.46	**31.27**	
M48	5	21	**80.77**	**21.96**	
M49	1	25	96.15	54.00	55.15
M49	1	25	96.15	49.19	
M49	0	26	100.00	62.27	
M50	0	26	100.00	81.85	79.99
M50	0	26	100.00	75.54	
M50	0	26	100.00	82.58	
M51	0	26	100.00	75.31	77.17
M51	0	26	100.00	79.54	
M51	0	26	100.00	76.65	
M52	0	26	100.00	67.15	64.09
M52	0	26	100.00	64.31	
M52	1	25	96.15	60.81	
M53	0	26	100.00	83.38	84.92
M53	0	26	100.00	83.42	

	Female	Male	Proportion	Gender rating	Overall rating
M53	0	26	100.00	*87.96*	
M54	0	26	100.00	84.54	*85.95*
M54	0	26	100.00	85.04	
M54	0	26	100.00	*88.27*	
M55	0	26	100.00	84.31	84.94
M55	0	26	100.00	82.27	
M55	0	26	100.00	*88.23*	
M56	1	25	96.15	76.81	82.94
M56	0	26	100.00	83.50	
M56	0	26	100.00	*88.50*	
M57	0	26	100.00	59.27	65.03
M57	0	26	100.00	65.69	
M57	0	26	100.00	70.12	
M58	0	26	100.00	60.77	64.06
M58	0	26	100.00	69.19	
M58	0	26	100.00	62.23	
Mean			96.75	64.48	
SD			8.37	21.75	
T59	26	0	*100.00*	**-72.73**	**-64.06**
T59	23	3	*88.46*	**-64.73**	
T59	24	2	*92.31*	**-54.73**	
T60	9	17	34.62	7.65	13.68
T60	3	23	**11.54**	18.96	
T60	5	21	19.23	14.42	
T61	6	20	23.08	24.88	24.62
T61	3	23	**11.54**	23.88	
T61	1	25	**3.85**	25.08	

	Female	Male	Proportion	Gender rating	Overall rating
T62	3	23	**11.54**	*41.31*	20.12
T62	6	20	23.08	20.85	
T62	15	11	57.69	-1.81	
T63	22	4	*84.62*	-22.77	-24.81
T63	22	4	*84.62*	-25.77	
T63	21	5	80.77	-25.88	
Mean			48.46	-6.09	
SD			36.19	36.21	

APPENDIX 4: CORRELATIONS

	f_0	H1-H2	H1-A1	H1-A3	F_1 bw	% Jit	% Shim	NHR	VTI	SPI	CPP	CPPs	score
f_0	1												
H1-H2	.58**	1											
H1A1	.37**	.61**	1										
H1-A3	.35**	.47**	.72**	1									
F_1bw	.36**	.26*	.23	.20	1								
% Jit	-.08	.15	.11	.15	.04	1							
% Shim	-.41**	-.09	-.02	-.07	-.09	.41**	1						
NHR	-.19	-.13	-.40**	-.44**	.11	.31*	.32*	1					
VTI	-.21	-.09	-.13	-.31*	.04	.39**	.49**	.63**	1				
SPI	.12	.23	.28*	.64**	.05	.14	-.05	-.41**	-.32*	1			
CPP	.07	-.32*	-.32*	-.43**	-.06	-.42**	-.38**	-.09	-.23	-.35**	1		
CPPs	-.17	-.29*	-.34**	-.46**	-.08	-.29*	-.18	-.04	-.11	-.30*	.83**	1	
score	-.71**	-.19	-.14	-.16	-.38**	.22	.44**	.16	.17	-.07	-.34*	-.09	1

Table 4.1a: Correlations acoustic measurements and gender score for carrier phrase with *had.* Note: *significant at 0.05 level; **significant at 0.01 level

	Mean f_0	SD f_0	Semitone range	SD semitones	High freq	Low freq	score
Mean f_0	1						
SD f_0	.627**	1					
Semitone	.225	.671**	1				
SD semi	.260*	.882**	.735**	1			
High f	.753**	.803**	.699**	.545**	1		
Low f	.751**	.219	-.323**	-.171	.428**	1	
score	-.852**	-.540**	-.173	-.195	-.672**	-.677**	1

Table 4.1b: Correlations acoustic measurements and gender score for carrier phrase with *had.* Note: *significant at 0.05 level; **significant at 0.01 level

	f_0	H1-H2	H1-A1	H1-A3	F_1 bw	% Jit	% Shim	NHR	VTI	SPI	CPP	CPPs	score
f_0	1												
H1-H2	.54**	1											
H1A1	.36**	.53**	1										
H1-A3	.25*	.34**	.71**	1									
F_1bw	.30*	.28*	.46**	.17	1								
% Jit	-.21	-.18	-.23	.09	-.08	1							
% Shim	-.42**	-.30*	-.32**	-.08	-.18	.70**	1						
NHR	.01	.02	-.41**	-.31*	-.10	.44**	.42**	1					
VTI	-.25	-.09	-.23	-.37**	-.03	-.00	.31*	.28*	1				
SPI	.25*	.00	.36**	.60**	-.09	.09	-.14	-.20	-.50**	1			
CPP	.15	-.16	-.22	-.37**	.02	-.32*	-.17	.06	.18	-.26*	1		
CPPs	.02	-.08	-.19	-.29*	-.06	-.23	-.03	.09	.20	-.25*	.88**	1	
score	-.76**	-.21	-.31*	-.14	-.20	.31*	.45**	.10	.23	-.28*	-.38**	-.16	1

Table 4.2a: Correlations acoustic measurements and gender score for carrier phrase with *hodd.* Note: *significant at 0.05 level; **significant at 0.01 level

	Mean f_0	SD f_0	Semitone range	SD semitones	High freq	Low freq	Response
Mean f_0	1						
SD f_0	.540**	1					
Semitone	.254*	.738**	1				
SD semi	.216	.914**	.797**	1			
High f	.820**	.845**	.622**	.613**	1		
Low f	.685**	.104	-.401**	-.210	.449**	1	
Response	-.868**	-.462**	-.263*	-.176	-.701**	-.537**	1

Table 4.2b: Correlations acoustic measurements and gender score for carrier phrase with *hodd.* Note: *significant at 0.05 level; **significant at 0.01 level

	f_0	H1-H2	H1-A1	H1-A3	F_1 bw	% Jit	% Shim	NHR	VTI	SPI	CPP	CPPs	score
f_0	1.00												
H1-H2	.43**	1.00											
H1A1	.42**	.60**	1.00										
H1-A3	.18	.47**	.56**	1.00									
F_1bw	.38**	.43**	.46**	.36**	1.00								
% Jit	.20	.16	.10	.08	.55**	1.00							
% Shim	-.30*	-.02	-.13	-.06	.20	.52**	1.00						
NHR	-.32*	-.06	-.29*	-.23	-.01	.47**	.74**	1.00					
VTI	-.31*	-.22	-.05	-.17	.06	.10	.38**	.38**	1.00				
SPI	0.18	.29*	.15	.51**	.10	.06	-.13	-.18	-.42**	1.00			
CPP	-.20	-.25*	-.27*	-.41**	-.51**	-.66**	-.59**	-.34**	-.05	-.39**	1.00		
CPPs	-.61**	-.34**	-.36**	-.33**	-.49**	-.58**	-.37**	-.19	.01	-.23	.77**	1.00	
score	-.75**	-.42**	-.44**	-.30*	-.51**	-.27*	.06	.13	.00	-.01	.30*	.60**	1

Table 4.3: Correlations acoustic measurements and gender score for sustained vowel
Note: *significant at 0.05 level; **significant at 0.01 level