Exploring Relationships Between Communication Features, Gender Attribution Ratings, and Quality of Life for Transgender and Cisgender Communicators

by

Teresa Lily Doreen Hardy

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

REHABILITATION SCIENCE

Faculty of Rehabilitation Medicine
University of Alberta

Abstract

Background: Voice and communication modification training is a critical aspect of the gender affirmation process for many transgender people. Incongruence between communication characteristics and gender positioning can be a cause of gender dysphoria and lead to misattribution or being outed as a transgender person, which can have significant negative social consequences (e.g., discrimination, physical harm). Consequently, identifying the characteristics of communication that contribute most to conveying one's gender and masculinity-femininity is important for informing voice and communication modification training practices.

Objective: The two main objectives of my doctoral research were to 1) Identify a set of communication-based predictors (i.e., acoustic and nonverbal communication measures) of subjective ratings related to gender attribution; and 2) Explore relationships between communication characteristics and self-rated outcomes of femininity, communication satisfaction, and quality of life (QoL) for transgender women.

Methods: The objectives of my doctoral research were met through the completion of four studies. Data collection for all four studies occurred at one time, across two phases. In the first phase, communication and QoL data were collected from a group of transgender women and cisgender communicators (n = 40). The communicators' voices and gestures were captured during a cartoon retell task via simultaneous acoustic and motion-capture recordings. A unique constellation of 11 acoustic and six gestural variables subsequently were measured in 30-45 second samples of these recordings or in two standard speech tasks recorded during the same data collection session. In the second phase of data collection, a group of raters (n = 20) provided gender attribution and masculinity-femininity ratings for each communicator. These ratings were based on

samples of the cartoon description recordings presented in three modes: audio only (i.e., audio track), visual only (i.e., point-light display), and audiovisual (i.e., audio track and point-light display). Raters also rated vocal naturalness in the audio-only condition.

The first study identified a set of acoustic predictors of gender attribution, masculinity-femininity, and vocal naturalness ratings obtained in the audio-only presentation mode. The second study identified a set of communication-based (i.e., acoustic and gestural) predictors of masculinity-femininity ratings obtained in the audiovisual presentation mode. That study also explored differences in the perceptual ratings as a function of audio, visual, or audiovisual modes. The third study explored relationships between the set of communication variables and self-rated femininity, communication satisfaction, and QoL for the transgender women participants (n = 20). The fourth study investigated differences in motion-based nonverbal communication behaviors between groups based on gender attribution and gender positioning.

Results: Significant predictors of masculinity-femininity ratings in audio and audiovisual modes included speaking fundamental frequency (f_o) ($p < .001$; $p < .001$), average formant frequency ($p = .001$; $p = .006$), and sound pressure level (SPL) ($p = .001$; $p = .001$). F_o was the sole predictor of gender attribution ratings in the audio-only mode ($p = .047$), and f_o ($p = .002$), average formant frequency ($p = .001$), and rate of speech (p .022) were identified as significant predictors of vocal naturalness ratings. Masculinity-femininity ratings obtained in the audio-only mode were significantly more feminine than those made in the audiovisual mode ($p < .001$). Visual only mode masculinity-femininity ratings were not reliable. Three significant relationships were revealed between the communication variables and subjective ratings made by the transgender women

participants: 1) Use of palm-up hand gestures was negatively related to gestural femininity ratings (r_s = -.462, p = .040); 2) Use of palm-up hand gestures was negatively related to overall communication satisfaction (r_s = -.572, p = .008); and 3) Mean semitone range across utterances was positively related to overall QoL (r = .463, p = .040). Finally, cisgender women used vertical head movements (e.g., nodding) significantly more than transgender women [$F(2,36) = 5.06$, p = .012].

Conclusions: The results of my doctoral research advanced our understanding of the ways in which voice and communication characteristics contribute to gender presentations and attributions, and how they relate to subjective ratings of femininity, communication satisfaction, and QoL for transgender women. Together, these studies add to the growing evidence base informing voice and communication modification training and, ultimately, have the potential to positively impact quality of life and life participation outcomes for transgender and gender diverse people.

Preface

This thesis is an original work by Teresa L. D. Hardy. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "Talking the talk and walking the walk: Communication and quality of life in the male-to-female transgender population", Pro00060133, February 3, 2016. This research project also received operational approvals from Northern Alberta Clinical Trials and Research Centre (NACTRC) and Covenant Health Research Centre.

Chapter 2 of this thesis has been published as Hardy, T. L. D., Rieger, J. M., Wells, K., & Boliek, C. A. (2018). Acoustic predictors of gender attribution, masculinity-femininity, and vocal naturalness ratings amongst transgender and cisgender speakers. Journal of Voice, p. 1–16. Advance online publication. http://doi.org/10.1016/j.jvoice.2018.10.002 I was responsible for study design with mentorship; ethics and institutional operational applications with mentorship; data collection and analysis; manuscript preparation, and overseeing the submission and review process. I also presented the study findings at the Fall Voice Conference, 2018, Seattle, WA. C. A. Boliek was the supervisory author and was involved with study design and manuscript edits. J. M. Rieger was involved with study design and manuscript edits. K. Wells assisted with participant recruitment and manuscript edits.

Acknowledgements

I have been so fortunate to have the support of a great many people throughout the course of my doctoral program and for that, I am eternally grateful.

First and foremost, I would like to thank my direct supervisors, Carol and Jana. You have both been amazing mentors and have taught me so much about how to be a good researcher, writer, and teacher. I am so grateful for the opportunities you have provided me; your guidance, support (scientific and emotional), resources, and networks; and the countless hours you have spent helping me (and pushing me!) to be successful. Your individual and collective brilliance continues to astound me.

I'd also like to thank Kris Wells for serving on my supervisory committee. Kris, you have been instrumental in ensuring that I (and my research) was appropriately informed by historical context and current trends in LGBTQ2S research and advocacy. I have greatly benefitted from your connections to key players in gender affirmative care and in the community. I look forward to continuing to collaborate on initiatives that will improve service-provision for transgender/gender diverse people in Alberta.

Data collection and analysis would not have been possible without the expert assistance of individuals from the Glenrose Rehabilitation Hospital, Communication Sciences and Disorders, Linguistics, Computing Sciences, and the private sector. Justin Lewicke and Daniel Aalto, your knowledge of motion capture, measurement/analysis, and computer coding was invaluable for measuring nonverbal communication in my doctoral research. Dr. Ben Tucker and Matthew Kelley, I greatly appreciate your advice in acoustics and for your continual support in measuring my sizeable set of acoustic

variables. Thanks also go out to the Computing Science students at the University of Alberta (Shawn Anderson, Dylan Cassidy, Ashley Fegan, and Mickael Zerihoun) who created an earlier iteration of the Gender Finder program.

My immense gratitude also goes out to my friends and family who have been there for me continuously over the past six and a half years, to celebrate my successes, encourage me when the mountain just seemed too high, and even to tell me to get a grip when needed. In particular, I'd like to thank Andréa Benoit May, Janine Chesworth, Leslie Chivers, Gabi Constantinescu, Gillian Lourie, Ashley McKillop, and Rainie Sharpe. Thanks also to Bruichladdich for providing me with sanity-saving pet therapy and making sure I got up from my desk to move every day. Thanks also go out to my Mom, of course, for being there to chat when I needed it and for being my cheerleader. It makes me proud when you say you're proud of me.

I was extremely fortunate to have received multiple awards and fellowships over the course of my doctoral program, which made the research possible, allowed me to reduce my clinical hours for three years, and helped me travel to learn and present my research outcomes. Thank you to the Social Sciences and Humanities Research Council of Canada, Alberta Innovates, University of Alberta, Alberta College of Speech-Language Pathologists and Audiologists, Faculty of Rehabilitation Medicine, Alberta Health Services, and the Graduate Students Association. I would not have been able to complete my doctoral research or program without your financial and in-kind support.

Finally, my sincerest thanks and appreciation go out to the individuals who volunteered their time to participate in my research. It truly would not have been possible without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abbreviations

A, V, AV Mode	Audio, Visual, and Audiovisual mode
AXI Head	Total axial head movement
COR Head	Total coronal head movement
dB	Decibel
Elbow%	Mean percent elbow:shoulder width
DME-WM	Direct Magnitude Estimation Scale – Without Modulus
F_0	Speaking fundamental frequency
F1, F2, F3	First, second, and third vowel formant
GAS	Gender Affirmation Surgery
GD	Gender diverse
Hand Mvt	Total hand movement
Hz	Hertz
LGBTQ2S	Lesbian, Gay, Bisexual, Transgender, Queer, Two-Spirit
MS Head	Total midsagittal head movement
NHR	Noise-Harmonic Ratio
Palmup%	Percent of time the lead hand was in the palm-up position
PLD	Point-Light Display
QLI	Quality of Life Index
QoL	Quality of Life
RoS	Rate of Speech
SLP	Speech-Language Pathologist
SPI	Soft Phonation Index

SPL	Sound Pressure Level
ST	Semitone
TG	Transgender
TVQ^{MtF}	Transsexual Voice Questionnaire (Male-to-Female)
VFF	Vowel Formant Frequency
VRQoL	Voice-Related Quality of Life

Glossary of Terms

Cisgender	An adjective used to describe individuals for whom gender positioning is congruent with sex assigned at birth.
Gender	The social constructs of being a man/masculine or woman/feminine, involving one's subjective sense of self, behaviors, physical presentation, and roles as well as and how they are attributed by others.
Gender Affirmation Process	The process of making social, medical, and/or legal changes to increase congruence between gender positioning, presentation, and/or desired attribution as determined by each individual. May also be referred to as transition.
Gender Affirmation Surgery	Surgical interventions used to better align the body or physical appearance with gender positioning and desired gender presentation. May also be referred to as Sex Reassignment Surgery, Gender Confirmation Surgery, bottom surgery, or top surgery.
Gender Attribution	The way others perceive one's gender. May also be referred to as perceived gender.
Gender Diverse	An umbrella term used to describe people who position or present their gender in ways that differ from socio-cultural norms. This group of people can include individuals whose gender positioning is non-binary.
Gender Performativity	The ways one experiences and produces or expresses their gender as influenced by socially constructed expectations and norms.
Gender Positioning	One's subjective sense of their gender. May also be referred to as gender identity.
Gender Presentation/Expression	The ways in which one expresses their gender. Gender may be outwardly expressed via behaviors, appearance (e.g., hair style, clothing), communication, name, pronouns, etc.
Intersex	An adjective used to describe individuals whose anatomical sex characteristics do not fit the conventional classifications of female or male.

Non-Binary	A term used to describe people who position or express their gender along a continuum, beyond the categories of man or woman.
Point-Light Display	A constellation of moving dots that be joined by lines that is created from motion capture data to represent human movement.
Sex	Medical classification of a person as male or female based on biological, physiological, and anatomical characteristics such as chromosomes, hormones, and primary and secondary sex characteristics.
Transgender	An adjective used to describe individuals for whom gender positioning and/or presentation is not congruent with sex assigned at birth.
Transgender Man/Transmasculine Person	A person who was assigned female at birth but identifies as a man/boy or positions gender in a masculine fashion.
Transgender Woman/Transfeminine Person	A person who was assigned male at birth but identifies as a woman/girl or positions gender in a feminine fashion.
T-Unit	A unit of syntactic complexity composed of one main clause plus its associated subordinate clauses.
Voice and Communication Modification Training	Training provided by speech-language pathologists as part of the gender affirmation process that guides clients in changing communication behaviors to be better aligned with expressed gender positioning, desires for gender presentation, and/or gender attribution.

CHAPTER 1: INTRODUCTION

Background

"I was born in the wrong body." This trope is one commonly used by transgender (TG) individuals to describe their lived experience (Prosser, 1998). Their gender positioning,[1] that is, who or what they feel they are inside, as far as where they fall on the continuum between man and woman (Azul & Hancock, 2018), is not congruent with their physical bodies. The terminology in this relatively new area of study, often referred to as transgender studies, is frequently evolving (E. Coleman et al., 2011), at times variable (Samons, 2009), and often contested and debated. There are, however, several key terms that provide clarity regarding different experiences of gender diversity.

It is beneficial to begin an introduction on the topic of transgender studies with an explanation of the difference between sex and gender. This very differentiation is fiercely debated; however, for the purpose of this dissertation work, and in keeping with scientific and research literature, they are defined as follows. Sex refers to the medical classification of a person as male or female based on biological, physiological, and anatomical characteristics such as chromosomes, hormones, and primary and secondary sex characteristics (Beemyn & Rankin, 2011). In contrast, gender refers to the social constructs of being a man/masculine or woman/feminine, involving one's subjective sense of self, behaviors, physical presentation, and roles as well as and how they are attributed by others (Beemyn & Rankin, 2011). Gender positioning, first mentioned in the

[1] Terminology is continually evolving in this area: what was previously referred to as gender identity is now starting to be referred to as gender positioning. Gender positioning will be used for the majority of the dissertation; however, at times it will be used interchangeably with gender identity, especially in discussions of background and theory in which identity was the accepted term and conceptualization.

introductory paragraph, is further defined by Samons (2009) in reference to gender identity as an innate sense of self, the way "one perceives oneself with regard to gender, and how the continuum of masculinity and femininity is experienced by self, rather than how one is perceived and experienced by others" (p. 22). This self-concept is thought to develop very early in life (Laqueur, 1990) and may or may not be consistent with a person's sex assigned at birth (Beemyn & Rankin, 2011).

The gender normative view that prevails in modern Western society assumes sex and gender to be congruent; however, that is not the case for many people. Individuals who are gender diverse (GD) are those who position or present their gender in ways that differ from socio-cultural norms, representing a spectrum of gender possibilities that may be stable or fluid over time (The 519 & Rainbow Health Ontario, 2016)[2]. This umbrella term encompasses transgender individuals and those whose gender positioning is non-binary. Non-binary is a term used to describe people who position or express their gender along a continuum, beyond the categories of man or woman (The 519 & Rainbow Health Ontario, 2016; Richards et al., 2016). The term 'transgender' also is considered to be an umbrella term, representing a diverse group of people.

Transgender individuals may further describe themselves or be described as transgender women/transfeminine or transgender men/transmasculine. A transgender woman is a person who was assigned male at birth but identifies as a woman/girl,

[2] A media reference guide was used as a reference for certain terminology used in the dissertation. This guide was created based on the work of community-based organizations, advocacy groups, LGBTQ2S service-providers, as well as by Canadian governmental human rights policy. The terminology presented in the guide was deemed to reflect current inclusive language recommendations and explicitly defined terms that routinely are used as part of the gender- and sexual minority-specific vernacular.

whereas a transgender man is a person who was assigned female at birth but identifies as a man/boy. Transgender or gender diverse (TG/GD) people may or may not hold binary views of gender.

It is difficult to determine how many people identify as transgender given the limitations of census and survey practices; however, prevalence rates have been suggested to range from 1:11,900 – 1:45,000 for transfeminine individuals and 1:30,000 – 1:200,000 for transmasculine individuals (De Cuypere et al., 2007). These rates are considered to be minimum estimates (E. Coleman et al., 2011) and new sampling procedures suggest the prevalence to be approximately 1:200 (Conron & Scott, 2012) or 0.6% (Flores, Herman, Gates, & Brown, 2016). These numbers may appear to be small, especially when compared to other groups accessing healthcare services; however, the transgender population is one that is particularly vulnerable, as will be discussed in detail in upcoming sections. This vulnerability supersedes, to some extent, the size of the population.

History of Transgender Identity

The first references to transgender identities did not occur in Western cultures until the late 1800s when sexologists such as Richard Freiherr von Krafft-Ebing described sexual inversion – a condition equivalent to modern descriptions of transgenderism – in the context of homosexuality (Prosser, 1998). Gender diverse identities, however, had long existed across the globe in groups such as the Hijras of India, the Mahu of Tahiti, Shamans of various Eurasian countries, and even Two-Spirit First Nations peoples from North America (Brown & Rounsley, 2003; Samons, 2009). Advancements in theory, related technology, and advocacy have since occurred with each subsequent decade.

The early twentieth century witnessed the first gender affirmation surgeries (GAS) in Europe (Prosser, 1998) and the first use of 'gender' as a term differentiated from sex (Gherovici, 2010). It was mid-century, however, that brought some of the most pivotal changes. D. O. Cauldwell coined the term 'transsexual' in 1949 (Brown & Rounsley, 2003; Prosser, 1998; Shelley, 2008) and that same year Alfred Kinsey, the famous sexologist, referred the first transgender cases to Dr. Harry Benjamin (endocrinologist and sexologist) for treatment in the United States (Gherovici, 2010). Within a few years, an American ex-G.I. named Christine Jorgensen publicly revealed her successful transition from male to female via hormones and surgical procedures performed in Denmark. This publicity vastly increased public awareness, altering opinions of transgender individuals and gender affirmation surgery, and ushering in "a new era of sex change in the U.S." (Gherovici, 2010, p. 88). Dr. Harry Benjamin, often known as the father of transgender research, took over Ms. Jorgensen's medical and psychological care and went on to publish "The Transsexual Phenomenon" (1966), the foundational text that distinguished between transsexuals and cross-dressers, and outlined guidelines for care related to the gender affirmation process (Shelley, 2008). Gender affirmation, previously known as transition, involves making social, medical, and/or legal changes to increase congruence between gender positioning, presentation, and/or desired attribution. The release of these first informal 'Standards of Care' coincided with the opening of the gender identity clinics at Johns Hopkins Hospital in the U.S. and the Clarke Institute in Canada (Shelley, 2008). Although these clinics purported to assist transgender individuals with the gender affirmation process, very few individuals were approved for

surgery at that time (Shelley, 2008). Moreover, with the availability of endocrine and surgical interventions came the necessity for a diagnosis and diagnostic criteria (Stone, 2006), which led to the medicalization of transgender identities (Shelley, 2008).

Gender Identity Disorder (GID), the psychopathology attributed to gender diversity, first appeared in the Diagnostic and Statistical Manual (DSM) in 1980 (Stone, 2006). This diagnosis remained in the DSM for another three decades. The 1990s brought the 'Transsexual Revolution' and advocacy for transgender rights (Prosser, 1998; Samons, 2009) and has been likened to the gay movement of the 1970s (Gherovici, 2010). Public awareness of transgender issues became more widespread at that time thanks, in part, to popular media and the advent of the Internet. These advances created a society characterized by increased tolerance and appreciation of diversity: Transgender individuals were perceived to be less of a threat to existing cultural ideologies (Samons, 2009). Members of the general public were not the only ones to benefit from the knowledge imparted by those involved in the transgender movement. Healthcare professionals also gained a better understanding of what it meant to be a transgender person. As a result, GID was replaced by Gender Dysphoria in the DSM-5, making the distinction that the diagnosable psychological condition was the *distress* associated with the discrepancy between gender identity (positioning) and biological sex and not the discrepancy itself (American Psychiatric Association, 2013). Being transgender was no longer considered pathological in and of itself. Consequently, the major goals of affirmation-related care became focused on mitigating feelings of distress rather than 'curing' people of being transgender.

Prior to providing affirmation-related care, it is important for health care professionals, including speech language pathologists, to understand transgenderism and appreciate prevailing theories of transgender identity development. This understanding may increase cultural competence and sensitivity as well as assist the clinician in creating appropriate training goals (in conjunction with the client) and approaching training in a way that is based on contemporary theory. Present-day theories have developed from and have been influenced by a diverse group of sociological, psychological, and historical perspectives including, but not limited to psychoanalysis, post-structuralism, post-modernism, feminism, and queer theory. These perspectives have contributed to current beliefs about relationships between gender, sex, and sexuality; gender identity development; the nature of gender (i.e., binary vs. existing as a continuum) and transgenderism (i.e., an example of hysteria vs. an example of 'normal' existence); the representativeness of gender norms; and the mechanism(s) underlying the desire for physical and social gender affirmative changes. Several of the related ideologies also have served as catalysts for the modern transgender movement. While interesting, it is beyond the scope of this paper to explore these theories at length; therefore, only the prominent contemporary theories that may drive clinical practice service will be discussed.

Sociological Perspectives and Theories of Gender Identity

Contributions of queer theory. Queer theory is a relatively new theoretical orientation based on a post-structuralist view of gender and sexuality that aims to reject gender normativity and heteronormativity (Nagoshi & Brzuzy, 2010; Valocchi, 2005). It attempts to rethink possibilities for the relationships between gender, sex, and sexuality; has adopted the view that gender and sexuality are performative or socially constructed;

and rejects the notion that certain sexual identities are associated with particular behaviors or desires (Beemyn & Rankin, 2011; Nagoshi & Brzuzy, 2010; Valocchi, 2005). Queer theory contends that current taxonomies are inadequate to describe all the feelings, experiences, and behaviors of gender and sexuality and, therefore, allows for and even welcomes the integration of 'opposing' identities (e.g., a woman with a beard) (Valocchi, 2005). Expected relationships are purposely inverted or reversed (Prosser, 1998) so as to break the cycle of performative behaviors and undermine existing norms. This process has been referred to as 'queering'. What is unique to queer theory is the belief that identities or positioning are not fixed but can be fluid over time (Butler, 2004; Gherovici, 2010).

Queer theory fits particularly well for transgender individuals who do not identify with the gender binary, seeing themselves as somewhere in between these two, some would argue, artificial end points. Within this perspective, a person can position (and present themselves) at any point in the gender continuum and this point need not remain fixed. For example, one day an individual may position their gender closer to the 'woman' end of the spectrum and present in a more feminine fashion while on another day, that person may identify and present in a more masculine way. Having an understanding of culture-specific expectations or assumptions for gendered behaviors may help an individual to live in a way that feels authentic and congruent with their gender positioning.

'Queering' of gender also accommodates the transitional period at the beginning of the gender affirmation process when a TG/GD person may not fit within societal gender expectations (Prosser, 1998). Prior to commencing the gender affirmation process, many transgender women have lived fully as men and expressed their gender in masculine

ways. Some may even have exceptionally masculine physiques (e.g., very muscular, very tall, a lot of body hair) or pursue stereotypically masculine life roles (e.g., armed forces, father), activities, or behaviors in an effort to suppress or alter their authentic gender positioning. These aspects of presentation cannot be changed instantly, nor even very quickly in some cases; therefore, it may become necessary for these individuals to start living in their affirmed gender role before they fully look, socially function, or behave in ways that society recognizes as being characteristic of a woman. Moreover, some may not wish to present their gender in a normative manner. Removing the expectation that people must conform to gender-appropriate behaviors, feelings, or physical characteristics creates a 'safe space', so to speak, for TG/GD people as they attempt to navigate through their gender affirmation process.

Queer theory is limited in its ability to fully represent the diversity of transgender people, however, in that it does not allow for embodiment of the binary (i.e., identifying as a woman or a man), effectively denouncing the lived experience of those who do position their gender in that way (Prosser, 1998). Many TG/GD people seek to align their sex and gender and to conform to societal expectations for gendered behaviors, which is beyond what is possible in queer theory (Prosser, 1998). The role of the physical body in identity formation and in the gender affirmation process underscores another limitation of the theory. Like post-structuralist theory and some feminist theory, queer theory suggests that the body is immaterial in identity development since gender and sex are independent of each other and culturally constructed (Prosser, 1998). This view cannot account for the

sense of congruency and completeness that many TG/GD individuals feel after undergoing gender affirmation sex surgery (GAS) (Prosser, 1998). For these reasons, queer theory and transgender theory can be allied but cannot be unified (Prosser, 1998).

Transgender theory. Transgender studies is an emerging field that is tasked with the challenge of developing a theory that adequately represents the lived experience of people who identify as transgender (Shelley, 2008). Assumptions of unity amongst the members of any social group overlook the intersecting factors (e.g., race, socioeconomic class) that also contribute to the experience of each individual (Butler, 1990; Nagoshi & Brzuzy, 2010). This critique is especially pertinent to transgender theory because of the fundamental differences that exist between those who identify with the binary and those who do not (Shelley, 2008). Whereas non-binary individuals embrace the post-structural rejection of gender binaries and the fluidity of gender espoused by queer theory, for instance, TG/GD people who identify with the binary take a modernist stance, embracing the gender binary and rejecting fluidity of gender over time (Shelley, 2008). A theory that is too extreme in either direction threatens to take away the political voice and sense of community of one group or the other (Shelley, 2008). The common ground between the two views, and where a cohesive transgender theory should perhaps focus, is on the promotion of a different view of the relationship between gender and sex: one that is not deterministic and allows for movement along the continuum, be it to the 'ends' or to a point in between (Whittle, 2006).

Most, if not all, of the sociological perspectives and psychological theories that have contributed to the understanding of transgenderism have historically neglected the role of the body in the development of gender identity (Prosser, 1998). Transgender theory,

however, stresses the importance of this relationship, purporting that gender can be physically embodied (Nagoshi & Brzuzy, 2010; Prosser, 1998). This view accounts for the experience of incongruence that is felt prior to gender affirmation for some TG/GD people and the sense of cohesiveness that is realized afterward. It also validates the narrative, or described experience, of being born in the wrong body, one that is shared by many transgender people.

The importance of engaging in narrative is characteristic of the transgender experience. Autobiographical narrative is the mechanism by which many transgender people discover and come to terms with their gender positioning, finding themselves through the telling of their story. Indeed, Prosser (1998) suggested, "transsexual autobiography produces gender identity" (p. 120). Narrative also is the mechanism by which transgender individuals are able to access medical interventions and transition to become their authentic selves (Prosser, 1998): It is the story, rather than the body, that designates one as TG/GD. Historically, there were certain episodes that psychiatrists listened for in a person's narrative to indicate whether or not that individual was truly transgender (since there were no objective criteria on which to base a diagnosis) (Stone, 2006). Queer theorists criticized this practice, alleging that such 'gate-keeping' led to reinforcement of gender norms and limitations in gender expression as a consequence of people saying what needed to be heard in order to get the diagnosis and, ultimately, the interventions they wanted (Spade, 2006).

A final, and very important benefit of narrative is that it allows TG/GD individuals to share their stories, educate others, and advocate for social justice. Narrative has played a key role in the transgender movement, which is said to have started in response to the

discrimination and repudiation experienced at the hands of radical cultural feminists, dispelling myths about what it means to be a TG/GD person (Prosser, 1998). Stone (2006), whose retort to Janice Raymond's scathing tome (Raymond, 1979) is credited to be at the heart of the movement, suggested embracing a "posttranssexual" narrative. She argued that people must tell their stories and identify themselves to the world for who they are so that conceptions may change, making identity diversity possible (Stone, 2006).

As a movement, transgender studies share common ground with intersex, queer, gay, and feminist groups (Butler, 2004). Intersex and transgender groups both advocate for the freedom to determine or state their gender and the queer group joins them in the opposition of identity legislation (Butler, 2004). Feminist and LGBTIQ (Lesbian, Gay, Bisexual, Transgender, Intersex, and Queer) groups all lobby for a "set of values in which gender loses some of its power of oppression, in which separate and distinct voices are not only heard but also listened to, and in which a better set of values is followed" (Whittle, 2006, p. 202).

Transgender theory is eclectic, drawing influence from fields as diverse as psychoanalysis, post-modernism, post-structuralism, feminism, and queer theory (Nagoshi & Brzuzy, 2010; Shelley, 2008). It is not, however, fully defined by these other theories, placing value in areas where other theories do not. Nagoshi (2010) succinctly described transgender theory as follows:

> Transgender theory encompasses and transcends feminist and queer theory by explicitly incorporating ideas of the fluidly embodied, socially constructed, and self-constructed aspects of social identity, along with the dynamic interaction and integration of these aspects of identity within the narratives of lived experiences. (p. 432)

Various studies and published autobiographical narratives have provided valuable information for the creation of theories relating to the development of transgender identities. They have revealed that most individuals have an awareness of being different at a very young age, come to recognize incongruences between sex and gender in their early teens, and begin to identify as transgender a couple of years later (Beemyn & Rankin, 2011). Several models have been proposed to describe the trajectory of transgender identity development in general and, more recently, for various gender diverse identities, specifically (e.g., transgender women). These models vary in terms of how many stages they contain, but most include a period of anxiety or confusion about gender positioning, a period when individuals identify primarily as transgender, a period when they deny or reject this transgender identity, and a stage when gender expression matches gender positioning (Beemyn & Rankin, 2011). Models with greater numbers of steps tend to include more detail surrounding physical transformation or embodiment of gender positioning and disclosure of gender positioning to family, friends, and others. In general, these models mirror the gender affirmation process facilitated by physicians, psychiatrists, and other health care professionals such as speech-language pathologists.

According to the most recent version of the Standards of Care for the Health of Transsexual, Transgender, and Gender-Nonconforming People (E. Coleman et al., 2011), when individuals present to gender clinics they must obtain a diagnosis of gender dysphoria from a qualified psychiatrist. Then, they will be referred to an endocrinologist to commence hormone treatment and start to work toward transitioning to living full-time as their affirmed gender. There currently is a requirement that people live full-time for at least one year, referred to as the 'Real Life Experience', prior to being approved for surgical interventions so that they are able to experience what it would be like to live in that particular gender role and, therefore, make an informed decision regarding the pursuit of these essentially irreversible changes. If they do wish to proceed with surgery, a second letter of referral from another qualified mental health professional (e.g., psychologist) is required. Not all individuals decide to pursue GAS, which has become more acceptable in transgender communities in recent years (Beemyn & Rankin, 2011). Many consider the process just described to be outdated and community members and healthcare professionals increasingly are advocating for the adoption of an informed-consent model. This type of model allows each individual to determine the services and interventions they require to affirm their gender and to access those services via their primary care provider.

Beemyn & Rankin (2011) proposed a different schema for transgender identity development based on their groundbreaking study of transgender identity and lived experience that surveyed and interviewed over 3000 transgender individuals from all over the United States. Unlike the other models of identity development, Beemyn & Rankin's schema does not suggest a linear progression through a series of stages. Instead, a number

of milestones (i.e., common themes) are provided that are specific to each type of gender diverse identity. The schema allows for individuality of experience in terms of which milestones are experienced and the way they are experienced. For transgender women, the milestones include acceptance and expressing the transgender identity, repressing the identity secondary to negative reactions from others, gaining knowledge and meeting other transgender people, identifying as a transgender person, accepting oneself as a woman, undergoing physical affirmation with hormones and/or surgery, deciding whether or not to disclose to others, and becoming comfortable with oneself even if unable to have their gender correctly attributed.

Unfortunately, correct gender attribution or not being identified as a TG/GD person can be a major determinant of safety. The current societal climate often is discriminatory toward transgender individuals and frequently leads to harassment and even violence (Bauer & Scheim, 2015; Beemyn & Rankin, 2011; Grant et al., 2011; James et al., 2016b). Societal views are difficult to change, especially for such fundamental constructs as sex and gender (Brown & Rounsley, 2003); therefore, regardless of one's personal opinion regarding the existence of such categories, they remain very much related to a person's socio-cultural belonging (Prosser, 1998). The prejudicial behaviors of others also can lead to mental health problems (e.g., depression or anxiety) for the TG/GD person and may contribute to the negative emotions associated with gender dysphoria, having dire consequences for the individual. For example, in a national survey conducted in the United States, the rate of attempted suicide was found to be an astounding 41% amongst transgender individuals as compared to only 1.6% in the general population (Grant et al., 2011). Similar results of attempted suicide were reported for Trans-

Ontarians with 43% of respondents of the TransPulse Survey reporting they had attempted suicide in their lifetime compared to 3.82% in the general population (Bauer & Scheim, 2015). Correct attribution becomes important in these situations because the more 'visible' a difference is, the greater the chance the transgender individual becomes a target for mistreatment or has a negative experience (Shelley, 2008).

Communication Perspectives

The challenge in creating a gender presentation that is congruent with gender positioning often is the catalyst that prompts TG/GD individuals to seek the services of speech-language pathologists (SLPs) for voice and communication modification training. The vast majority of clients accessing such services are transgender women, who wish to feminize their communication behaviors; however, the number of transmasculine clients accessing services is beginning to grow and indicates an important area for future study. Modifications to voice and mannerisms represent a critical aspect of the gender affirmation process for many people because the way one communicates has a considerable impact on the way one's gender is attributed by others and, consequently, how one is accepted or able to participate in society (Byrne, 2007; Pasricha, Dacakis, & Oates, 2008). Moreover, incongruence between communication behaviors and gender positioning can be a major source of dysphoria. Finally, when communication behaviors are incongruent with what is expected based on attributed gender, the individual may be read as a TG/GD person and be at risk for discrimination, harassment, and violence discussed previously, which can limit their ability to fully participate in life (Bauer & Scheim, 2015; Beemyn & Rankin, 2011; James et al., 2016b).

Whereas exogenous androgen hormone treatments often result in voice changes for transmasculine individuals (E. Coleman et al., 2011), feminizing hormone treatments (i.e., anti-androgens and estrogens) do not significantly change the voice for transgender women (Gelfer & Schofield, 2000; Wolfe, Ratusnik, Smith, & Northrop, 1990). As Samons (2009) so astutely commented, "testosterone is a one-way street" (p. 11). When young males are exposed to increased levels of endogenous androgen during puberty, their bones, cartilages, and muscles become bigger and bulkier than those of their pre-pubescent or female peers (Krause, 2006). The structures of the vocal tract are no exception; post-pubertal males have larger vocal tracts (particularly in the structures of the larynx) than females (Klatt & Klatt, 1990; Oates & Dacakis, 1997). Gender-related differences in voice are a product of these anatomical (and associated aerodynamic) differences (Holmberg, Oates, Dacakis, & Grant, 2010). For example, men have deeper voices and different resonance characteristics than women (Hillenbrand & Clark, 2009). Transgender women who undergo gender affirmation after puberty will have experienced the same changes to their voices and vocal tracts secondary to androgen exposure. For this reason, communication changes must primarily be made behaviorally.

Although voice is known to be an important gender marker (Dacakis, Davies, Oates, Douglas, & Johnston, 2013; Owen & Hancock, 2010; Van Borsel, De Cuypere, & Van den Berghe, 2001; Wolfe et al., 1990), several other communication variables are thought to contribute to gender attribution, including those related to speech, language and nonverbal communication (Adler, Hirsch, & Mordaunt, 2006; Dacakis, Oates, & Douglas, 2012; Holmberg et al., 2010; Oates & Dacakis, 1983; Van Borsel et al., 2001).

Having an understanding of gender diversity not only makes clinicians more culturally competent but also allows for client-centered service provision. For example, understanding whether an individual identifies as binary or non-binary can guide the selection of training targets. For example, transgender women who identify with a gender binary may wish to present in a manner that aligns more closely with normative feminine behaviors, whereas non-binary individuals may aim for a more gender-ambiguous presentation. Regardless of gender positioning, each client's desires for gender presentation ultimately guide goal selection. In order to set goals that are specific, measureable, and appropriate, and that will help the client achieve their desired gender presentation, it is important to understand the characteristics of communicators who are attributed as men, women, or ambiguous in gender. Furthermore, we must understand which of these factors are most salient to gender attribution as well as to the perception of femininity and vocal naturalness as these are factors related to 'authentic' gender expression.

Acoustic measures of voice. Gender-related differences in voice can be objectively quantified using a variety of acoustic measures. Acoustic differences in the voices of cisgender men and women, as well as individuals attributed as men and women (transgender or cisgender), are thought to exist in measures such as speaking fundamental frequency (f_o), vocal tract resonance characteristics and vowel formants, intonation (e.g., upward intonation shifts), rate of speech (RoS), sound pressure level (SPL), and voice quality. These parameters have been investigated to varying extents in the cisgender and

transgender literature. The relative salience of each measure as a cue to attributed gender remains unclear in most cases. It is, therefore, valuable to review the existing literature in each of these areas individually.

Speaking fundamental frequency (f_o). It has been well established in both transgender and cisgender literature that speaking f_o plays an important, even critical, role in attribution of gender (Dacakis et al., 2013; Gelfer & Mikos, 2005; Gelfer & Tice, 2013; Hillenbrand & Clark, 2009; Holmberg et al., 2010; Leung, Oates, & Pang Chan, 2018; E. J. M. McNeill, Wilson, Clark, & Deakin, 2008; Oates & Dacakis, 1997; Van Borsel et al., 2001; Wolfe et al., 1990). In fact, a recent meta-analyses reported this voice feature explained 41.6% of the variability in listener attribution ratings (Leung et al., 2018). Furthermore, f_o has been positively related to femininity ratings (Carew, Dacakis, & Oates, 2007; Gelfer & Tice, 2013; Gorham-Rowan & Morris, 2006; Hardy et al., 2016; Owen & Hancock, 2010) and was found to account for 78% of the variability in listeners' perception of femininity in one study (Owen & Hancock, 2010). What remains unclear, however, are the frequency ranges that characterize speakers who are consistently attributed as a particular gender and how f_o interacts with other communication variables in measures of gender attribution.

Threshold values of 145-160 Hz initially were suggested for an individual to be attributed as a woman (Oates & Dacakis, 1983; Spencer, 1988; Wolfe et al., 1990); however, later studies did not support this value as sufficient (Gelfer & Schofield, 2000) and suggested it was higher than 180 Hz (King, Brown, & McCrea, 2012). Other studies suggested that communication factors besides frequency contribute to gender attribution, especially when the speaker's f_o falls within a 'gender ambiguous' or 'overlapping' range

in which a person may be attributed as either a man or a woman (Gelfer & Bennett, 2013; Gelfer & Tice, 2013; Van Borsel et al., 2001). This 'gender ambiguous' range has been defined previously as 150-185 Hz (Mordaunt, 2006) and falls in the middle of the normative f_o ranges for males and females, reported as 220 +/- 20 Hz and 120 +/- 20 Hz, respectively (Andrews, 1999).

Many communication-based cues have been proposed, both clinically and in the research literature, as contributing to gender attribution and/or perceived femininity such as vocal tract resonance, intonation patterns, sound pressure level, voice quality, rate of speech, and body language (Adler, Hirsch, & Mordaunt, 2012; Dacakis et al., 2012; Gelfer & Bennett, 2013; Holmberg et al., 2010; Leung et al., 2018; Oates & Dacakis, 2015; Owen & Hancock, 2010; Van Borsel et al., 2001; Wolfe et al., 1990); however, with the possible exception of vocal tract resonance, their relative contribution to gender attribution is not well understood.

Vocal tract resonance and vowel formants. Vocal tract resonance is the sympathetic vibration of air in various cavities of the vocal tract (e.g., oral, nasal, pharyngeal) in response to vibration of the vocal folds (Baken & Orlikoff, 2000) and relates to timbre, or tone quality, of the voice (King et al., 2012). In the communication literature, vocal tract resonance typically is represented by vowel formant frequencies. Each vowel formant refers to "a single frequency at which vocal tract transmission is more efficient than at nearby frequencies" (Baken & Orlikoff, 2000, p. 27). The first formant (F1) corresponds with tongue height or mouth opening: As tongue height increases or the mouth closes, the frequency of F1 decreases. The second formant (F2) corresponds with the forward or back location of the point of restriction in the resonating cavity (i.e., front or back tongue

placement in the production of vowels): The frequency of F2 is higher for vowels articulated at the front of the mouth (e.g., /i/), and lower for those articulated at the back of the mouth (e.g., /u/). The third formant (F3) does not clearly correspond with an articulatory movement; however, decreased degrees of lip rounding are known to raise all of the first three vowel formants (Hixon, Weismer, & Hoit, 2008).

Vocal tract resonance has long been suspected to play a key role in distinguishing the voices of men and women (Gelfer & Bennett, 2013; Gelfer & Tice, 2013; Günzburger, 1995; Hillenbrand & Clark, 2009; Mount & Salmon, 1988). This suspicion is supported by the finding that female vowel formants are significantly higher [by approximately 20% (R. Coleman, 1983)] than those of males (Gelfer & Bennett, 2013; Hillenbrand, Getty, Clark, & Wheeler, 1995). The investigation of vocal tract resonance characteristics and vowel formants has gained increased attention in transgender research over the past 15 years (Carew et al., 2007; Gelfer & Mikos, 2005; Gelfer & Schofield, 2000; Gelfer & Tice, 2013; King et al., 2012). Although the resonant quality of the voice (as measured by vowel formants) now is generally considered to contribute to gender attribution and/or perceived femininity (Dacakis et al., 2012; Davies, Papp, & Antoni, 2015; Gelfer & Bennett, 2013; Leung et al., 2018; Oates & Dacakis, 2015), the strength of the contribution and the salience of each vowel formant remains unclear. For example, studies have shown relationships between auditory-perceptual ratings of gender and/or femininity and each vowel formant (F1, F2, F3) in isolation (Carew et al., 2007; Gelfer & Bennett, 2013; Gelfer & Tice, 2013; Deborah Günzburger, 1993, 1995; Mount & Salmon, 1988) as well as between those ratings and an average of the three formants together (Gelfer & Bennett, 2013). To complicate matters, some studies have reported

that none of the first three vowel formants were significantly related to masculinity-femininity ratings (Gelfer & Schofield, 2000; Hardy et al., 2016), nor did they significantly contribute to attributed gender in transgender women speakers (Gelfer & Mikos, 2005; Gelfer & Schofield, 2000; King et al., 2012). Study limitations such as small sample sizes, differences in training history, and stage of gender affirmation may have limited the ability to find group differences in these cases. Moreover, the study completed by Gelfer and Mikos (2005) only looked at vowel formants in the context of isolated vowels, a context rarely encountered in daily interactions. These results were not replicated when vowel formants were extracted from connected speech (Gelfer & Bennett, 2013).

Gelfer and Bennett (2013) provided compelling evidence in support of the contribution of vowel formants in their study of cisgender speakers. These authors measured the accuracy of gender attribution of cisgender men and women speakers when naturally produced vowel formant frequencies were paired with f_os that were artificially manipulated to be gender appropriate, characteristic of the other gender in the binary pairing, or fell within the 'gender ambiguous' range. They found that 63.5% of the variance in listener judgment of gender in connected speech could be accounted for by F1. The controlled manipulation of f_o in the context of connected speech and purposeful inclusion of gender ambiguous frequencies coupled with a moderate sample size (n=15) made this work important.

Another study, completed by the present author and colleagues (Hardy et al., 2016) found that F2 positively related to the perception of vocal naturalness (i.e., typically-sounding or authentic voice) in transgender women (n=25) who had not yet undergone

communication feminization training. That is, as F2 increased, vocal naturalness ratings also increased. These results are unexpected given that the vast majority of speakers were attributed as men and would be expected to have low vowel formants. It is possible that speakers articulated in a way that raised F2, which they perceived as feminine. As such, they may have felt more comfortable with their communication, which was conveyed to the listeners as natural speech. Vocal naturalness is clinically relevant because most people accessing services for communication feminization, or any other communication difficulty (e.g., stuttering, dysarthria) for that matter, are seeking to sound like what they consider to be a typical speakers.

Intonation. Another aspect of communication that commonly is believed to differ between men and women and contribute to 'authentic' or natural-sounding speech is the intonation pattern or melody of speech. In fact, Andrews and Schmidt (1997) touted, "intonation is the most well-documented supra-segmental speech marker that differs between masculine and feminine voice patterns" (p. 308). Intonation also is one of the most commonly targeted goals in communication feminization treatment (Hancock & Garabedian, 2013), speaking to the importance put on this parameter by both clinicians and clients.

In general, it is believed that frequent use upward intonation patterns ('upspeak') and greater pitch variability when speaking contribute to listener attribution of a speaker as a woman whereas monotone or falling intonation patterns and variability in loudness for emphasis contribute to listener attribution of a speaker as a man (Gelfer & Young, 1997; Hancock, Colton, & Douglas, 2014; Leung et al., 2018; Mordaunt, 2006; Oates & Dacakis, 1997; Wolfe et al., 1990). Despite the widely held belief that intonation patterns

are important to gender attribution, investigation of this communication parameter is relatively lacking in the transgender literature and limited to only a few studies.

Wolfe et al. (1990) investigated the intonation patterns of 20 transgender women speakers. These speakers were grouped according to attributed gender, based on short conversational utterances. Unlike many other studies, speakers were grouped nearly equally: Nine in the attributed as a woman group and 11 in the attributed as a man group. These authors found that speakers who were attributed as men used downward intonations that were significantly greater in extent than those used by speakers attributed as women. Intonations were defined as a change in frequency by five Hz or more without a stoppage of phonation. Speakers attributed as men also used significantly more level intonations than those attributed as women (i.e., a change of less than five Hz) whereas speakers attributed as women used significantly more upward intonations. Significant differences also were reported in use of intonation shifts, which were defined by those authors as a change of at least five Hz between the end of one utterance and the beginning of the next. Speakers that were attributed as women used significantly more downward shifts and fewer level shifts than those attributed as men. Significant correlations between these same measures and ratings of femininity-masculinity revealed similar relationships.

The results of Wolfe et al.'s research support the belief that feminine speakers use more upward intonation patterns and greater variability in intonation patterns. A major limitation of the study, however, was in the operational definitions of 'intonations' and 'shifts': The five Hz cutoff is substantially smaller than the two semitone (ST) cutoff used to define intonations and shifts in other studies (Gelfer & Schofield, 2000; Hancock,

Colton, & Douglas, 2014). For example, two semitones can represent a difference of more than 200 Hz - 400 Hz (Baken & Orlikoff, 2000). Such minute changes in frequency may not be clinically relevant. Furthermore, the semitone scale is considered to be more appropriate than the hertz scale for making comparisons between auditory perception and acoustic measures because of its nonlinearity: Humans are thought to perceive sound in a similar, nonlinear, fashion (Henton, 1989).

Gelfer and Schofield (2000) investigated similar intonation measures as did Wolfe et al.; however, Gelfer and Schofield defined an intonation shift as a "change in frequency, with or without interruption in phonation, of at least 2 semitones" (p. 26). Intonation measures included upper and lower f_o limit (Hz), f_o range (ST), and extent and number of upward and downward intonation shifts. Speakers (n=13) were grouped according to attributed gender (man or woman) and compared on all acoustic measures. Correlations also were completed between these measures and ratings of femininity-masculinity. The authors found the upper limit f_o to be the only measure related to intonation that differed significantly between attributed gender groups. Furthermore, this measure was the only one significantly correlated with femininity-masculinity ratings: As upper limit increased (mean of approximately 300 Hz), so too did rated femininity. The study was limited by the small sample size of transgender women speakers, especially for the group of individuals attributed as women (n=3), which may have resulted in a failure to demonstrate differences that truly existed between the groups.

Finally, Hancock, Colton, and Douglas (2014) conducted a study investigating gender differences in intonation patterns during a picture description task. Comparisons were made between gender groups (i.e., cisgender men, cisgender women, transgender women,

and transgender men), attributed gender groups (i.e., man and woman), and between transgender women who were consistently attributed as women, consistently misgendered, and whose gender was attributed as ambiguous. Each group was compared on a variety of intonation measures including: duration of utterance; mean, minimum, maximum, starting, and ending frequency; ST range; ST slope; and percentage of upward and downward intonation shifts. Intonation shifts were defined as in Gelfer and Schofield (2000). These measures also were correlated with femininity measures completed by naïve listeners. The results of the study indicated that there were no significant differences between gender groups for any intonation measure; however, there were significant differences (with medium-large effect sizes) between attributed gender groups, which, arguably, is most relevant to the goals of voice and communication modification training. Individuals attributed as women used a greater ST range and more upward intonation shifts than those attributed as men. Although not statistically significant, the same trend was observed in transgender women speakers who consistently were attributed as women: They used more upward intonation shifts and a greater ST range than those who were not. Individuals attributed as women used a mean of 37% upward intonation shifts and a ST range of 5.56 ST compared to 21% and 4.26 ST in speakers attributed as men, respectively. Another interesting finding was that transgender women speakers who consistently were misgendered used a greater percentage of downward intonation shifts than even cisgender men speakers. None of these acoustic measures of intonation correlated significantly with femininity ratings. The authors concluded that intonation measures were weak predictors of perceived femininity.

The results of these studies were partially in keeping with those of Wolfe et al. (1990) in terms of differences in intonation shifts and ST range between attributed gender groups; however, these measures were not found to be significantly correlated with measures of masculinity-femininity. Moreover, another study investigating relationships between acoustic measures and speaker- and listener-perceived femininity ratings reported opposite findings with respect to ST range: Femininity ratings were significantly related to a narrow ST range (Owen & Hancock, 2010). The authors reported that 26% of the variability in speaker-rated femininity and 36% of the variability in listener-rated femininity could be accounted for on the basis of ST range.

Finally, the results of a recent systematic review found that intonation measures such as frequency of upward and downward intonation patterns, ST range, and f_o variability were associated with listener attributions of a speaker gender (Leung et al., 2018). Of note, the systematic review included studies of both cisgender and transgender participants.

Taken together, these studies suggest that intonation measures may have some influence on gender attributions and perceived femininity, particularly measures such as ST range and percentage of upward intonation shifts. Further research is needed to confirm previous results and determine the relative salience of intonation as a cue to gender, especially when f_o is within the gender ambiguous range. Furthermore, Gelfer and Bennett (2013) suggested that more research is needed in the area of prosody, specifically with respect to "rate, rhythm, and 'melody'" (p.11) in order to better understand the role of these features as auditory-perceptual cues to gender in connected speech.

Rate of Speech. The cisgender literature is somewhat inconclusive with respect to gender differences in rate of RoS. Some studies have found that speech rate is faster in men than women in both reading and conversational contexts (Fitzsimons, Sheahan, & Staunton, 2001; Verhoeven, De Pauw, & Kloots, 2004), whereas others report RoS to be similar between gender groups (Tsao & Weismer, 1997; Tsao, Weismer, & Iqbal, 2006; Van Borsel & De Maesschalck, 2008). Often, popular belief is that women speak at a faster rate than men. This perception, as suggested by Michelle Mordaunt during an interview with Amy Norton (2000), may be due to the tendency for women to speak in a pattern of "short bursts followed by a pause". Based on this description, one might hypothesize that women have a slower speech rate [i.e., the number of syllables or words per unit of time, including pauses (Tsao & Weismer, 1997)] but a faster articulation rate [i.e., the number of syllables or words per unit of time, excluding pauses (Tsao & Weismer, 1997)] than men. This hypothesis was not supported, however, by Verhoeven, de Pauw, and Kloots (2004), who found that Dutch men had speech and articulation rates that were 6% faster than those of Dutch women, on average. Similarly, Bouw (as cited in Gunzburger, 1989) reported that women took longer than men to read single words when speaking normally and when whispering.

There are relatively few data with respect to RoS in transgender women speakers and the data that do exist reflect the same inconsistencies seen in the cisgender literature. For example, Günzburger (1995; 1993; 1989) found that transgender women took significantly longer to read a given word or sentence when speaking in 'woman mode' versus 'man mode', suggesting those speakers believed a slower rate of speech sounded more feminine. Similar findings were reported in a study that found men who were

perceived to sound effeminate had a slower RoS (in words per minute) than those perceived to have a 'typical' masculine voice (Terango, 1966). Conversely, Van Borsel and Maesschalck (2008) found no differences in RoS between cisgender men, cisgender women, and transgender women speakers when reading a standardized passage. Of note, neither of the studies involving transgender women speakers investigated differences in RoS as a function of gender attribution or perceived masculinity-femininity, which are more relevant to communicative success than gender positioning. Moreover, both studies used reading tasks and did not investigate the role of RoS in a conversational context. In their review of current research findings in MtF communication, Dacakis, Oates, and Douglas (2012) proposed that further research is needed to investigate RoS as a potential gender marker.

Sound Pressure Level. According to Gelfer and Young (1997), SPL (perceptual correlate = loudness) is one of the most salient perceptual features of the voice. Furthermore, SPL has been shown to differ in small but significant ways between cisgender speakers. For example, in a conversational context, men speak two to three dBSPL louder than women (Boonin, 2006a; Gelfer & Young, 1997) and have a wider conversational loudness range, which relates to use of greater sound pressures in their inflections (Gelfer & Young, 1997). A recent systematic review investigating acoustic contributions to listener perception of speaker gender also found this measure to be associated with gender attribution ratings (Leung et al., 2018). Although this parameter has a presence in clinical discussions of communication features to be considered in

voice and communication modification training (Adler et al., 2012; Dacakis et al., 2012; Oates & Dacakis, 1983, 1997, 2015), very little research has been done to investigate the role of SPL in gender attribution, especially as it relates to transgender speakers.

Early studies by Günzburger (1995; 1989; 1993) demonstrated that transgender women (n=6) used significantly lower SPL when speaking in their feminine voices and wider SPL ranges when speaking in their baseline masculine voices. Unfortunately, this study did not take into account listener perceptions of the voice and whether or not loudness differences would result in differences in gender or femininity ratings.

Holmberg et al. (2010) measured vocal SPL in a larger group (n=25) of transgender women speakers during different speaking tasks including typical conversation. Naïve listeners subsequently rated the gender of these speakers based on the verbal samples. The results of the study showed that gender attribution was based solely on loudness for some of the participants (n=4). Half of these participants were attributed as women and presented with softer voices and the other half were attributed as men and presented with louder voices. The authors concluded that "lower SPL and increased use of low voice intensities may help contribute to a more sucessful female voice…" (p. 520). Since these early studies point to a link between vocal loudness and gender attribution, it follows that we should understand how this parameter contributes to gender attribution for a larger group of transgender women speakers.

Voice quality. Voice quality is another communication construct that has been proposed to differ between men and women. Breathiness is one of the most commonly used descriptors of feminine voice quality and is thought to be the most important feature of voice quality to address in voice and communication feminization training (Dacakis,

2002). Auditory-percpetual ratings of breathiness, tenseness, lightness and avoidance of glottal fry have been associated with listener attributions of a speaker as a woman (Leung et al., 2018). Conversely, perceived roughness has been associated with more masculine-sounding voices (Hardy et al., 2016; Holmberg et al., 2010).

Cisgender studies have reported that women's voices are perceived to be breathy (King et al., 2012; Klatt & Klatt, 1990) and are more breathy than men's voices (Klatt & Klatt, 1990; Mendoza, Valencia, Muñoz, & Trujillo, 1996; Södersten, Hertegård, & Hammarberg, 1995). This breathiness is thought to be caused by the friction of air passing through a posterior 'chink' in the glottis during phonation in women (Klatt & Klatt, 1990; Södersten et al., 1995), which results, acoustically, in "greater amounts of aspiration noise in the higher formant regions in female voices" (Oates & Dacakis, 1997, p.179).

One study compared normal and breathy productions of the vowel /α/, as in 'father', with respect to perceived femininity in a group of cisgender women speakers (Van Borsel et al., 2009). The results demonstrated that breathy samples were judged to be significantly more feminine. This perception was the same whether normal and breathy samples were presented randomly or in matched pairs. Similarly, Andrews & Schmidt (1997) investigated the acoustic and perceptual characteristics of the voices of 11 male cross-dressers. Voice samples were produced in masculine and feminine modes and subsequently were judged by a cohort of 88 listeners. Listeners judged voices produced in the feminine mode to be more breathy than those produced in the masculine mode overall and the scale of breathiness was found to be useful in discriminating between the two modes.

Very few studies have compared sexes or genders on acoustic measures of voice quality (e.g., jitter, shimmer) (Oates & Dacakis, 1983), especially with respect to transgender speakers. There is no singular acoustic measure to represent this perceptual construct, perhaps in part because voice quality can encompass vastly different physiological states. For example, the configuration of the larynx while producing a tense voice may be different than when producing a breathy one. Nevertheless, there are several measures that have been used in previous research to measure these subjective qualities. For example, measures of cycle-to-cycle variability in vocal fold timing (i.e., jitter, RAP) and amplitude (i.e., shimmer) have been related to perceived vocal harshness or hoarseness (King et al., 2012; Wendahl, 1966a, 1966b). Furthermore, Bhuta and Garnett (2004) correlated subjective measures of voice quality using the Grade, Roughness, Breathiness, Asthenia, Strain (GRBAS) scale, a commonly-used clinical assessment tool, with instrumental measures of vocal noise taken with the Multi-Dimensional Voice Program, an acoustic measurement tool that is part of the Multispeech/Visiptich acoustic analysis suite (KayPentax, 2008). These authors found ratings of roughness to correlate significantly with Noise-Harmonic Ratio (NHR) only and ratings of breathiness to signficantly correlate with Soft Phonation Index (SPI) only. Van Borsel, Janssens, and De Bodt (2009) cautioned that the correlation between acoustic measures of breathiness and perceived breathiness is generally low (p. 293).

Within the transgender literature, Holmberg et al. (2010) found that ratings of breathiness (as made by SLPs) were not significantly related to any other variable they studied, including gender attribution. Self-rated vocal fry (“croaky voice”) as rated by

transgender women speakers, however, was significantly negatively correlated to gender ratings, where higher ratings were representative of very female speakers. These results suggested rougher or 'croakier' voices were associated with male voices.

Owen & Hancock (2010) found only weak correlations between self-perceptions of femininity obtained from transgender women participants and Noise to Harmonics Ratio (NHR), jitter percent, and shimmer percent in their study. The results were similar with listener perceptions of femininity except that shimmer percent was found to be moderately correlated with perceptions of femininity. The authors suggested that a lack of variation in vocal qualities and breathiness may have limited the contribution of voice quality to gender perception. All participants in their study had perturbation values that were within normal limits. King et al. (2012) had similar normal findings for jitter and shimmer amongst the 21 transgender women speakers and nine cisgender men and women speakers in their study.

Hardy et al. (2016) reported that the mean jitter (measured as Relative Average Perturbation – RAP) in the pre-training voices of 25 transgender women speakers was elevated when compared to the norms for both cisgender men and women. The authors speculated that these voices, which were nearly all attributed as the voices of men, sounded rough or breathy in quality to the listeners who provided the gender ratings. That same study reported that shimmer percent was found to be negatively related to the perceived vocal naturalness of these participant's voices. That is, as shimmer percent increased, the naturalness of the speakers' voices decreased. Since virtually all speakers were attributed as men, it was suggested that shimmer percent may have represented a feminine voice quality such as breathiness.

Finally, a systematic review completed by Leung et al. (2018) found SPI to be an acoustic measure of voice quality that significantly contributes to gender attribution. That review also provided evidence that other acoustic variables of voice quality such as frequency perturbation, amplitude perturbation, and aperiodicity do not significantly contribute to gender attribution.

Overall, the evidence is inconclusive; however it appears that certain acoustic measures of voice quality may contribute to subjective ratings of gender attribution, femininity-masculinity, and/or vocal naturalness (at least for unfamiliar listeners) and additional research is needed to more clearly describe this contribution. Given previous research results, SPI, NHR, and shimmer may be the acoustic measures most important to investigate due to their associations with relevant voice qualities and/or subjective ratings.

Non-verbal communication.

Evidence from cisgender research. In addition to the relatively well-documented differences in verbal communication, differences also have been shown to exist between cisgender men and women in terms of nonverbal communication. Study findings in this area have been replicated and confirmed over decades of research (LaFrance, Hecht, & Paluck, 2003; Vrugt & Kerkstra, 1984). These contrasting behaviors involve amount of interpersonal space used during communicative interactions, body posture, gazing, facial expression (especially smiling), and use of touch (Hall, 2006; Kennedy & Camden, 1984). Nodding (Hall & Friedman, 1999; Helweg-Larsen, Cunningham, Carrico, & Pergram, 2004) and use of hand gestures (Hall & Friedman, 1999; Hall, 2006; Hostetter & Hopkins, 2002; Nicoladis, Pika, Yin, & Marentette, 2007) also have been studied.

Gender or sex-based differences in nonverbal communication have been shown to exist across age groups (Hall & Friedman, 1999), contexts (Hall & Friedman, 1999), and cultures (Yang, 2010), although the majority of research has focused on university-aged students in North America or Europe. These studies suggest women are more expressive than men and better at using and understanding nonverbal communication (Briton & Hall, 1995; Hall & Friedman, 1999; Hall, 2006; Hostetter & Hopkins, 2002).

Unlike verbal differences between gender groups, which are thought to be less pronounced in recent years, nonverbal behaviors appear to be more resistant to change, preserving the group distinction (Kennedy & Camden, 1984). Goss (1982, in Kennedy & Camden, 1984) suggested that the maintenance of stereotypical behaviors may be due to the less conscious nature of nonverbal compared to verbal communication. LaFrance (1981, as cited in Epstein, 1986) provided three additional hypotheses to explain the continued differences in nonverbal communication behaviors between genders. The first suggests that characteristic gestures and movements function as a means of gender expression. As a result, an individual is able to present their gender positioning to others, with whatever degree of masculinity or femininity they wish, by using nonverbal behaviors that are considered to be more characteristic of a particular gender. The second hypothesis suggests that differences in nonverbal behaviors are the result of segregation of the sexes within certain contexts or cultures. This segregation effectively removes any mutual assimilation of behaviors that may occur between gender groups when they exist alongside one another. Third, differences in nonverbal communication may be the byproduct of social constructivism: Members of a community or culture are taught that

girls and boys or men and women behave in certain ways. For example, it's "acceptable in Western culture, even encouraged, for girls and women to be emotionally expressive" (Briton & Hall, 1995, p. 81).

One of the most popular theories, historically speaking, was that of Henley (1977), which suggested that social status was responsible for disparities in nonverbal communication. At that time, Henley proposed that women demonstrate more submissive behaviors that reflect their assumed inferior status in relation to men. More recent literature has shown that these gender differences are not due to status but, rather, socialization (Hall & Friedman, 1999; Hall, 2006). In fact, Hall and Friedman (1999) reported that gender differences were greater when the effects of status (i.e., the status within the company in which participants worked) were controlled. If these behaviors are indeed a result of socialization, they may be amenable to modification for the purposes of increasing congruency between gender positioning and gender presentation for transgender communicators. Furthermore, understanding how cisgender men and women differ with respect to their use of specific nonverbal behaviors may provide valuable guidance regarding which of these behaviors should be modified to achieve desired gender presentation characteristics.

Gestures. Kennedy and Camden (1984) operationally defined gesture as "any arm or hand movement which does not involve touching self or manipulating objects" (p. 97). This definition excludes behaviors such as scratching the face or twisting hair. Whereas cisgender men have been shown to have a tendency to move more in general when

communicating (Bente, Donaghy, & Suwelack, 1998), women use more meaningful gestures rather than fidgeting movements; that is, their movements correspond with the semantic content of their utterances (Hostetter & Hopkins, 2002).

Studies conducted in North America have shown that women use more hand gestures during story retell tasks compared to men (Hostetter & Hopkins, 2002; Nicoladis et al., 2007). The same gestural patterns were found for both monolingual (Hostetter & Hopkins, 2002) and second language English speakers (in both their first and second language) (Nicoladis et al., 2007). Nicoladis, Pika, Yin, and Marentette (2007) also found that women told longer stories than men and that the number of scenes retold and gesture rate (i.e., number of gestures used divided by number of word tokens spoken multiplied by 100) were positively correlated. That is, the more word tokens spoken, the more gestures were used. Consequently, these authors recommended controlling for task complexity (i.e., the number of scenes retold) when conducting research in this area.

The possibility also exists that men and women differ not only in amount of gestures used, but in type of gestures used as well. For example, studies have found women used more palm-up gestures when communicating, whereas men used more pointing behaviors (Peterson, 1976; Yang, 2010). Research focusing on differences in types of gestures used as opposed to frequency of movements is otherwise very limited. Bente, Donaghy, and Suwelack (1998) completed a study that included the coding of arm, shoulder, and hand movements (in addition to head movements), giving them the potential to examine differences in types of gesture used by men and women. Unfortunately, all movements in that study were pooled together to create measures of 'activity' (i.e., frequency, duration, and temporal distribution of movements) and complexity of behaviors. The authors were

able to provide further evidence that men and women differed in terms of frequency of movement (i.e., men moved more in general, especially when interacting with a familiar partner, and women's movements were positively correlated with their amount of speaking) but were not able to provide information about gender differences in types of gestures used amongst their participants.

Not all research has demonstrated the aforementioned patterns of gesturing during communication. For example, amount of gesture used by men and women was reported in a study examining the gestures of Anglo-Australian and Chinese speakers (Jones, Gallois, Callan, & Barker, 1999). That study found that women used fewer communicative gestures than men. Taken together, evidence available in the current literature suggests there remains more to be learned about gender differences in use of gesture.

Head movements/Nodding. Similar to differences in use of arm and hand gestures, women are believed to use more nodding behaviors than men, at least in the context of listening to another speaker (Hall & Friedman, 1999; Helweg-Larsen et al., 2004). Unfortunately, the literature is lacking in information about gender differences in head movements used during speaking activities although there is evidence that such movements serve as cues to gender attribution when judged by naïve observers. Hill and Johnston (2001) found that rigid head movements (i.e., movements that do not involve change in shape like those of facial expression) such as nodding, shaking, or tilting, cued observers regarding the identity and gender of a given speaker. Those authors reported that the way a speaker moved or rotated the head significantly contributed to an

observer's ability to categorize the gender of an animated representation of a speaker correctly and suggested it was due to the idiosyncratic nature of those behaviors. Overall, however, this area of investigation remains vastly unexplored.

The majority of the nonverbal communication literature discussed thus far has been based on studies using the coding of video recordings of speakers when interacting with conversational partners. This type of measurement allows for the description of differences in gesture in terms of frequency (e.g., number of movements observed) or type of movements used. Another body of research exists, however, that utilizes movement-tracking techniques that provide information about the characteristics of the behaviors directly.

Biological motion. Biological motion analysis is used as a means to measure human (or animal) movement. The earliest systems, first used in research in the early 1970s, attached light sources to major joints on the body and tracked them during movement. Contemporary systems use passive markers (also attached to major body joints), which are tracked by a sophisticated camera/computer system that uses the movement information to generate point-light displays (PLD). A PLD appears as a constellation of moving dots. Typical marker locations include the head shoulders, elbows, wrists, hips, knees, and ankles (Mather & Murdoch, 1994; Wöllner & Deconinck, 2013). An additional marker placed at the top of the sternum also may be used as a fixed point of reference (Wöllner & Deconinck, 2013).

Perceptual research has confirmed that as long as there are a sufficient number of points in the display, they are seen as human once they begin to move (Wöllner & Deconinck, 2013). Moreover, perceptual judgments based on PLDs are similar to those

that are made when observers can see the actual figure (Abernethy & Zawi, 2007). The benefit of using point-light displays, however, is that they eliminate the confounds associated with differences in body morphology (e.g., secondary sex characteristics, face) or personal style (e.g., clothing, hairstyle) (Johnson, McKay, & Pollick, 2011; Pollick, Kay, Heim, & Stringer, 2005) that are visible in regular video recordings or live observations.

In addition to recognizing that moving points represent humans, it generally is accepted that observers are able to accurately perceive the gender of a person based on a point-light display taken during a walking task (Kozlowski & Cutting, 1977; Mather & Murdoch, 1994). This ability is thought to be based on structural cues (i.e., cues related to body structure), motion cues (i.e., cues related to movement of body parts), or a combination of both and is most efficient in frontal or oblique views (Pollick et al., 2005).

Although there is less literature devoted to the investigation of motion as a cue to gender in comparison to body structure (e.g., the ratio of shoulder to hip size, shoulder breadth etc.), it is clear that certain types of motion provide cues for gender discrimination as well as emotion in point-light walkers (Johnson et al., 2011; Mather & Murdoch, 1994; Pollick et al., 2005; Troje, 2002a). These motion cues may even be more salient than cues related to body structure in gender discrimination tasks (Runeson & Frykholm, 1983). Examples of motion cues include lateral body sway (head and shoulder versus hip motion), cadence or rate of walking, stride length, arm swing, and velocity of movements (Mather & Murdoch, 1994). Of these, lateral body sway appears to be the most salient cue to gender in walking (Mather & Murdoch, 1994).

In addition to movements associated with locomotion, the existing biological motion literature provides information about non-locomotive arm and head movements, which may be more relevant to communication-based gender attribution.

Arm movements have been shown to effectively cue the identity and affect of a person to observers (Pollick, Lestou, Ryu, & Cho, 2002) in tasks such as throwing a ball (Johnson et al., 2011), conducting an orchestra (Wöllner & Deconinck, 2013), waving, knocking, and lifting (Pollick et al., 2002). These cues, however, appear to be more salient in some tasks than others. For example, Pollick, Lestou, Ryu, and Cho (2002) found that human observers were not able to reliably attribute gender from a point-light display of an actor who was waving, knocking, or lifting; however, a specialized computer algorithm was able to do so successfully. The authors concluded that the arm movements they studied did carry cues to gender but that human observers did not consistently use all the cues available to them when making gender judgments.

Wöllner and Deconinck (2013), however, found that human observers were able to recognize the gender of novice conductors from a point-light display taken during a conducting task and that judgments appeared to rely more heavily on motion cues than structure cues. It stands to reason that the findings may be similar in a conversational context given that both conversation and conducting tasks are more interactive than those studied in Pollick et al. (2002). That is, conducting and conversing both involve repeated instances of communication and responsiveness between the participants: A conductor and an orchestra respond to each other to make music and a speaker and listener take conversational turns. This type of repeated interaction does not occur with waving, knocking, or lifting.

Similar to arm movements, the biological motion literature has shown that movements of the head and face provide cues to gender independent of structure or surface cues (Hill & Johnston, 2001) as was discussed in the section on head movements and nodding. Taken together with evidence from the nonverbal communication literature, head movements and arm gestures appear to be worthy of further exploration.

In the review thus far, the biological motion literature suggests that the salience of gender cues as well as the ability for observers to use these cues is task- and body part-dependent (Pollick et al., 2002; Wöllner & Deconinck, 2013). Another factor to be considered is mode of stimulus presentation because it also may impact observers' ability to attribute gender.

Biological motion studies typically have used only the visual condition when assessing gender attribution or recognition. Wöllner and Deconinck (2013), however, assessed gender attribution of a group of conductors in audio, visual and audiovisual (AV) modalities. They found that these attributions were most accurate in the AV condition, but were not as reliable as typically is seen in gait studies. It appeared that the presence of auditory information (i.e., the orchestral music) influenced the gender attribution ratings in this study. Van der Zwan et al. (2009) found a similar interaction in their study of the impact of auditory information (i.e., the sound of foot strikes) on the gender attribution ratings made of ambiguously-gendered point-light walkers. When foot strikes were characteristic of a female walker, these ambiguous displays were perceived to be more female. The authors stated that, "the integration of information modalities

reduces the likelihood that the observer will be confronted with a scene that is perceptually ambiguous." (p. 373). For this reason, the AV modality may provide the best context for gender attribution ratings.

Evidence from transgender research. Despite the evidence that cisgender men and women differ in the way they communicate nonverbally and the evidence that these behaviors cue communication partners about the speaker's gender, there is a scarcity of research in the communication modification literature (Davies et al., 2015). The possible importance of nonverbal communication has not gone unnoticed in the clinical realm, however, as is evidenced by the fact that clinical training resources have devoted entire chapters to this domain (Hirsch & Boonin, 2012; Hirsch & Van Borsel, 2012). These chapters largely are based upon cisgender research; clinical experience; the general audience work of linguistics, communication, and body language experts; and other transgender or cross-dressing resources (e.g., available online). There is a dire need for evidence on which to base this practice. To date, only one study has addressed the influence of visually-perceived information during communication in the transgender population.

Van Borsel, De Cuypere, and Van den Berghe (2001) investigated observer perceptions of the voice and physical appearance of a group of transgender women. The authors presented the observers with recordings of these individuals reading standardized passages and the observers were asked to rate the 'femaleness' of the speaker. Three presentation modes were used in the study: visual only (i.e., video footage), audio only (i.e., voice recording), and audiovisual (i.e., recordings including both video and audio formats). The authors found a significant interaction between appearance and voice:

Speakers were perceived to be most female in the visual presentation mode, followed by audiovisual, and least female in the audio mode. A training focus on physical appearance consequently was recommended.

The generalizability of these results is questionable given that only 14 speakers were rated and the study has not yet been replicated. It also is possible that other transgender women speakers have more feminine voices than appearance, which would lead to different results. Moreover, the results of this study address group differences in physical appearance only and it is not clear on what information observers based their ratings. Perhaps observers were basing ratings on aesthetics (e.g., hair or clothing style), body morphology, or facial features or perhaps ratings were based on nonverbal communication. It also is possible ratings were based on a combination of all these factors. Without directly measuring certain factors, controlling for others, and including a comparison group (i.e., cisgender speakers), any conclusion in this regard would be conjecture only.

The body of literature on acoustic and non-verbal aspects of communication demonstrates that whereas these measures are believed to contribute to gender attribution and/or perceived masculinity-femininity or vocal naturalness, much remains to be learned with respect to the specific role and relative salience of each cue, especially when considering transgender communicators. For example, Hardy et al. (2016) reported that almost 50% of the variability in femininity ratings and more than 60% of the variability in vocal naturalness ratings remained unexplained by the nine acoustic variables measured in that study. The authors suggested that other parameters (e.g., RoS, prosodic characteristics) may be salient to such subjective measures. It also is important to gain an

understanding of how these communication variables relate to measures of quality of life, life participation, and communication satisfaction for transgender communicators if the ultimate goal is to identify voice and communication modification training targets. Such information is important because the ultimate goal of most SLP treatment or training programs is to improve the lives of the clients being served (American Speech-Language-Hearing Association, 2016).

Quality of life and communication satisfaction. For many transgender women, being attributed as a woman is related to certain aspects of life such as personal safety and acceptance in society. These individuals have described feeling a need to be read as a cisgender woman because of the negative opinions held about transgender people in western cultures and the negative emotions that accompany episodes of being misgendered (Byrne, 2007). The goals of being correctly attributed are multiple but related: to be seen by others in a way that is congruent with gender positioning and self-attribution; to avoid standing out as being 'different'; and to facilitate what Byrne (2007) dubs "getting on with life" (i.e., moving beyond being defined by transition and simply living as a 'regular' member of society) (p. 122).

Getting on with life involves being able to perform daily activities without limitations and to participate in society without restrictions. Unfortunately, herein lies the obstacle for many transgender women. Activities such as shopping, talking on the phone, using drive-through services, and fulfilling various life roles are challenging, often negative, experiences. The participants in Byrne's (2007) qualitative study documenting the lived experience of a group of transgender women, identified relationships (of all manner), employment, and community life as common areas of participation restriction they faced.

These restrictions are consistent with reported areas of discrimination experienced by TG/GD people in Canada and the United States (Bauer & Scheim, 2015; Grant et al., 2011; James et al., 2016b).

The ability to fulfill life roles and participate in society has obvious implications for a person's quality of life (QoL). For example, when a person is unable to participate in their life and community in the way or to the degree that they would like, they may experience feelings ranging from dissatisfaction to dysphoria or even depression. Since being correctly attributed is so intimately connected with a transgender woman's ability to participate in life roles (Byrne, 2007), it follows that increasing the likelihood of such attributions also may improve QoL for transgender women.

In her dissertation work, Byrne (2007) created a framework of describing attribution as a woman for transgender women showing that it is based on two factors: 1) what people see, and 2) what people hear. She further broke down the first factor (what people see) into how transgender women look and how they act. How transgender women look covered aspects such as surgery or aesthetics (e.g., clothing, hair style). While the results were interesting, they are beyond the scope of the proposed research. How transgender women act and how they sound, however, are directly in line with the proposed research aims of my dissertation work. Nonverbal communication behaviors fall under the former category and speech and voice fall under the latter. Behaviors that draw the attention of others or that are not consistent with the societal norm for behaviors expected for women negatively impact desired gender attribution according to Byrne's participants. Similarly,

voices that do not match the societal norm for women or are incongruent with the speaker's appearance in terms of degree of femininity may contribute to an individual being identified as TG/GD.

Using Byrne's framework, it is logical to hypothesize that communication behaviors are directly related to gender attribution outcomes and, therefore, impact QoL. Similarly, satisfaction with communication behaviors also may be related to QoL in that individuals may be more satisfied with their communication when it is congruent with and results in desired gender attributions and they are better able to participate in their communities. Byrne did not measure QoL directly in her research nor did she objectively measure the communication behaviors of her participants. Her framework was created using a grounded theory approach based on information gained from in-depth interviews. There is a paucity of published research investigating the relationship between communication behaviors and communication satisfaction or QoL in the TG/GD population.

Relationships between communication behaviors and communication satisfaction. The limited studies that have explored relationships between communication and such communicator-rated subjective outcomes have identified several factors that are related to communication satisfaction amongst transgender women. Satisfaction with voice, self-perceived pragmatic ability, and time spent in the gender affirmation process have significantly predicted communication satisfaction in this group (Byrne, Dacakis, & Douglas, 2003). Furthermore, in the same study, satisfaction with voice accounted for just over half of the 61% variance in communication satisfaction accounted for by the regression model, indicating it is important to consider in research involving this parameter. It is interesting to note, however, that only 16% of the variance in

communication satisfaction was predicted by average f_o, providing further evidence that other factors contribute to training success and client well-being. A limitation of the Byrne, Dacakis, and Douglas study was that only three variables were entered into the regression model as a result of the small sample size used (n=21). It is possible that other factors not included in their analysis may predict communication satisfaction for transgender women.

In another study, Pasrischa, Dacakis, and Oates (2008) identified five themes related to communication satisfaction amongst transgender women: 1) situations; 2) groups of people; 3) emotions; 4) other contributing factors (e.g., time of day, appearance etc.); and 5) features of communication. Of these themes, two are particularly relevant to the proposed line of research: features of communication and groups of people.

The participants in Pasricha, Dacakis, and Oates (2008) described five communication features that were related to communication satisfaction. These included voice, vocabulary, body language, speech, and intonation. These features represent multiple aspects of communication and further support studying relationships with communication factors other than f_o. In their study, participants indicated they were least satisfied with their voice and felt it was the aspect of their presentation that was most likely to lead to misgendering. They also reported their nonverbal communication was naturally feminine. These findings suggest links between both acoustic and nonverbal communication measures with perceptual outcomes. For example, a low f_o or vowel formant frequencies (VFF) may be related to negative outcomes whereas using more hand gestures may be related to positive ones.

In terms of groups of people, anecdotal reports and the available voice and communication feminization literature provide support for the idea that communication satisfaction varies for transgender women according to their communication partner (Pasricha et al., 2008). When partners are familiar and accepting (e.g., friends and family), communication satisfaction tends to be higher and not very dependent upon features of communication; however, when partners are strangers, maintaining feminine communication patterns becomes much more important (Pasricha et al., 2008). Being misgendered by strangers commonly is reported to be emotionally distressing to transgender women (Byrne, 2007). It is the frequency of this type of experience that voice and communication modification training aims to decrease in order to increase communication satisfaction and overall QoL. Since there is less distress associated with communicating with familiar partners, determining the predictors of correct gender attribution with unfamiliar partners is more valuable for informing voice and communication modification training. Moreover, Hancock, Krissinger, & Owen (2011) found observer (i.e., people other than the transgender communicator) perceptions of vocal femininity to be moderately positively correlated with QoL in transgender speakers. It is interesting to note, however, that the way transgender individuals rated the femininity of their own voices positively correlated even more strongly than observer perceptions, both with measures of QoL and communication satisfaction (Hancock et al., 2011; E. J. M. McNeill et al., 2008). This relationship speaks to the importance of measuring the perceptions of transgender speakers as well as, or instead of, observers. Such measures are conspicuously limited in the existing communication modification literature.

Taken together, the available research shows promising preliminary evidence for associations between communication measures and self-rated femininity and communication satisfaction. These relationships also may extend to QoL as several studies have shown significant correlations between communicator-rated femininity and/or voice satisfaction and voice-related QoL (VRQoL) (Dacakis, Oates, & Douglas, 2017a; Hancock et al., 2011; E. J. M. McNeill et al., 2008): higher self-ratings were associated with higher VRQoL. An introduction to the measurement of QoL will be discussed prior to reviewing relationships between these measures and communication behaviors.

Quality of Life Measurement. Service provision in Canadian health care, especially in the fields of rehabilitation medicine, currently emphasizes client-centered care. This means that treatment programs are directed by the needs and values of the client and that indicators of treatment or training success are based on client satisfaction and ability to participate in their daily lives. In the field of speech-language pathology (SLP), especially in the specialty area of voice, specific questionnaires have been used to measure the impact of the communication problem on the general voice client's life and have been used as indicators of quality of life and/or communication satisfaction (different, yet related, constructs). Tools such as the Voice Handicap Index (VHI) (Jacobson et al., 1997) and the Voice Related Quality of Life (VRQOL) (Hogikyan & Sethuraman, 1999) commonly are used with individuals experiencing voice disorders; however, these tools do not address critical aspects related to the communicative experiences of transgender

women. Moreover, communication satisfaction or quality of life tools should be population specific because each group has unique challenges and may value communication differently (Eadie et al., 2014; T'Sjoen et al., 2006).

Two tools have been developed specifically for transgender women: the Transsexual Voice Questionnaire for Male-to-Female Transsexuals (TVQ^{MtF}) (Dacakis & Davies, 2012) and the Functional Communication Satisfaction Questionnaire for Male-to-Female Transsexuals (FCSQ) (Dacakis, Pasricha, & Oates, 2005). Only the TVQ^{MtF} has published psychometric data and has been shown to have high content and construct validity as well as high test-retest reliability (Dacakis et al., 2013; Dacakis, Oates, & Douglas, 2017b, 2017c; Davies & Johnston, 2015). The TVQ^{MtF} was developed from an earlier tool, the Transgender Self-Evaluation Questionnaire (TSEQ), which was a modified version of the VHI created by Shelagh Davies. The TVQ^{MtF} was developed to measure a transgender woman's perception of her own voice and its effect on her ability to perform daily activities and participate in society (Dacakis et al., 2013), aspects related to both QOL and communication satisfaction.

In addition to being reliable, the TVQ^{MtF} also is in line with the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) (WHO, 2001), a framework designed to describe a person's state of health and resultant level of function as well as the interaction between these factors. Use of this framework is encouraged in health-related fields to create a common lexicon amongst practitioners and clients and to facilitate a comprehensive approach to service delivery. Application of the ICF to voice and communication modification training for TG/GD people is appropriate and of value (Hancock, 2017; Hardy, Boliek, Wells, & Rieger, 2013). Using this

framework guides clinicians to evaluate and address a client's ability to perform tasks (e.g., speaking), fulfill life roles (e.g., have meaningful relationships), and take part in society (e.g., hold a job), all of which may influence QoL. Taken together, psychometric data, population-specificity, and alignment with the ICF framework make the TVQ^{MtF} and ideal measure of VRQoL for transgender women.

Relationships between communication behaviors and Quality of Life. To date, only one study has explored relationships between communication variables and communicator-rated QoL with transgender women (Dacakis et al., 2017a). Those authors found no significant correlations between f_o, jitter, shimmer, and noise to harmonic ratio (NHR) and scores on the Transsexual Voice Questionnaire (TVQ^{MtF}) (Dacakis & Davies, 2012). The acoustic variables included in their study do not fully represent the features of communication identified as important to communication satisfaction in Pasricha et al. (2008) or those that are thought to cue gender attribution. Moreover, that study measured only VRQoL, offering a limited view of QoL as only one facet of a single aspect of a person's life (i.e., communication) was considered. QoL may comprise elements such as personal relationships, stress, and physical health, among others (Ferrans & Powers, 1985). As stated previously, communication, life participation, and personal safety are interrelated for many TG/GD people and all, therefore, may influence QoL. For this reason, a measure of overall QoL offers a more complete picture of a person's well-being than VRQoL. Indeed, Cruice, Worrall, and Hickson (2000) recommended using both generic and condition-specific QoL tools as outcome measures in SLP practice. To date, no studies have explored relationships between communication variables and overall QoL

amongst transgender women. Overall, relationships between communication behaviors and QoL remain underexplored and are worthy of further investigation (Dacakis et al., 2017a; Hancock et al., 2011).

In summary, a greater understanding of the communication factors that contribute to correct gender attribution, QoL, and communication satisfaction is needed. Furthermore, research investigating such relationships should take into consideration the specific, and potentially differing, perspectives of both unfamiliar observers and transgender women.

Perceptual ratings.

Gender attribution and masculinity-femininity. Given that attribution as the affirmed gender is deemed to be important by the majority of transgender women seeking voice and communication feminization training, it follows that communicative success should be measured, at least in part, through the subjective ratings of communication partners. Moreover, research focusing on communication patterns for the purpose of informing this training, also should include such measures.

Perceptual ratings typically involve attributions about the gender of the speaker and increasingly, ratings of masculinity-femininity. In the past, it has been common practice for studies to measure either attributed gender or perceived masculinity-femininity; however, it is perhaps most informative to measure both constructs as they provide related but distinct information. For example, gender attribution ratings indicate whether a communicator is perceived as a man or a woman; however, they do not provide further characterization of that communicator. For example, is the communicator perceived as a masculine or feminine man or woman? Similarly, masculinity-femininity ratings do not explicitly give information about gender attribution. For example, a speaker assigned the

highest femininity rating could have been attributed as either a very feminine woman or a very feminine man. Collecting both types of subjective measures provides a more complete assessment of the way a speaker is perceived.

Naturalness. Another quality that is important to address in both voice and communication modification training and research is how natural or, conversely, how unnatural a person communicates. Whereas it is possible to drastically change communication patterns to achieve desirable gender attribution or femininity-masculinity ratings, for many the ultimate goal is to be perceived as a typical speaker, to blend in with the masses, so to speak. If there is something odd or unnatural about the way a person communicates, it may draw undue attention and increase the liklihood of being recognized as a TG/GD person.

It is common practice in studies investigating communicative naturalness for researchers to refrain from providing a definition of the construct for participants (i.e., listeners) (Van Borsel & Eeckhout, 2008). Those studies that do define naturalness refer to speech that sounds normal to a listener and lacks qualities that may be distracting (Parrish, 1951; Ratcliff, Coughlin, & Lehman, 2002). This definition is appropriate for describing vocal naturalness as well.

Naturalness ratings have been used in communication research in a number of different areas such as stuttering, synthetic speech production, dysphonia, tracheoesophageal speech, and dysarthria (Bellaire, Yorkston, & Beukelman, 1986; Eadie & Doyle, 2002; Michael-Michelke, 1997; Nisltio, Tanaka, & Sakabibara, 2007; Ratcliff et al., 2002; Teshima, Langevin, Hagler, & Kully, 2010; Van Borsel & Eeckhout, 2008). These studies have found ratings to be related to general prosodic characteristics (Bellaire

et al., 1986; Nusbaum, Francis, & Henly, 1995); RoS (Nisltio et al., 2007; Ratcliff et al., 2002; Teshima et al., 2010; Van Borsel & Eeckhout, 2008); voice quality (i.e., breathiness, strain-roughness, clarity) (Michael-Michelke, 1997); monotonicity (Bellaire et al., 1986; Michael-Michelke, 1997); and fluency (Van Borsel & Eeckhout, 2008). Only one study, to the best of my knowledge, has attempted to measure communicative naturalness amongst transgender women communicators, specifically.

Hardy et al. (2016) attempted to determine acoustic predictors of perceived vocal naturalness amongst transgender women who had not received formal voice and communication feminization training from a SLP. Listeners were instructed to rate the voices based on whether or not something sounded out of the ordinary. Using step-wise multiple regression analysis, they found three models that significanly explained perceived vocal naturalness ratings made by naïve listeners. The first model identified only minimum frequency achieved during a maximum range task as being predictive of a more natural voice. Models two and three added the second vowel formant of /a/ and shimmer percent, respectively. The authors speculated that voices that did not extend to the lowest frequencies in connected speech or that were perceived to be rough or breathy in quality were rated as more unnatural. They also hypothesized that using certain articulatory postures that consequently raised F2 made the speakers feel more feminine, and thus, at ease, which was perceived by others to sound more natural.

A limitation of the Hardy et al. (2016) study was that intra-rater reliability was found to be only fair to good (ICC = 0.622-0.731) while inter-rater reliability was very good (ICC = 0.879). The authors suggested the low intra-rater reliability may have been due to training procedures and the measurement scale used. They recommended providing more

training with the rating scale and on the concept of vocal naturalness prior to commencing research trials. They also provided evidence from the literature that despite their widespread use, the type of scale used in the study [Visual Analog Scale (VAS)], has been found to be less reliable than other types of scales such as Equal Appearing Interval (EAI) or Direct Magnitude Estimation (DME) scales for a variety of communication variables. Researchers must design perceptual rating tasks carefully to ensure they are able to accurately and reliably address the aims of their research.

In summary, the existing literature demonstrates that multiple areas of communication contribute to gender attribution, perceived masculinity-femininity, and vocal naturalness. The evidence for the contribution of f_o and vowel formant frequcy to these subjective ratings is relatively strong whereas it is limited and/or inconclusive regarding the role of other measures such as SPL, intonation, voice quality, prosodic features, and nonverbal communication behaviors, especially when including TG/GD communicators. More research is needed to determine the relative contribution of these parameters, or combinations of these parameters, as predictors of observer ratings and to advance understanding of the relationships between communication features and communicator-rated subjective outcomes such as femininity-masculinity, communication satisfaction, and QoL.

Dissertation Objectives

The main objectives of this research were to describe the portfolio of communication factors that contribute to subjective ratings related to gender attribution as well as the relationships between communication factors and indicators of QoL and communication satisfaction for transgender women. Specific aims were to:

1. Determine acoustic predictors of observer-rated gender attribution (i.e., man, woman, ambiguous in gender), femininity, and vocal naturalness.
2. Determine communication-based (i.e., acoustic and nonverbal) predictors of observer-rated femininity.
3. Determine whether femininity ratings differ by stimulus presentation mode (i.e., audio, visual, audiovisual).
4. Determine whether upper body gestures differ between groups based on gender positioning and attributed gender.
5. Describe relationships between communication variables and communicator-rated femininity, QoL, and communication satisfaction.

Increasing understanding of the predictive relationships between communication variables and the identified subjective ratings will help inform voice and communication modification training so that goals are centered on the strongest cues to gender in conjunction with the areas identified as most important to the individual client. Moreover, advancing knowledge of these relationships will help to guide further refinement of gender presentation so that clients are able to communicate in a way that is more congruent with their specific gender positioning and desired gender presentation. These objectives are consistent with the principles of client-centered care now adopted by many service providers.

Structure of the Dissertation

The aims of the dissertation were addressed via a series of four studies. The first study identified acoustic predictors of observer ratings and the results are presented in Chapter 2 in a paper entitled "Acoustic Predictors of Gender Attribution, Masculinity-Femininity,

and Vocal Naturalness Ratings Amongst Transgender and Cisgender Speakers". Specific questions were: What acoustic measures predict: (1) gender attribution (as man, woman, or ambiguous in gender)?; (2) perceived masculinity-femininity?; and (3) perceived vocal naturalness? The first question was answered using logistic regression and the second and third questions were answered using standard multiple linear regression. Observer ratings collected in the audio only presentation mode served as the criterion variables and the set of acoustic measures served as the predictor variables. Study 1 (Chapter 2) was published in the *Journal of Voice* (Hardy, Rieger, Wells, & Boliek, 2018).

The second study expanded upon the first by including measures of nonverbal communication in the predictor set. That study identified communication-based predictors of observer-rated femininity and examined differences in femininity ratings between stimulus presentation modes, thereby addressing the second and third aims. Results are presented in Chapter 3 in a paper entitled "Contributions of Voice and Nonverbal Communication to Perceived Masculinity-Femininity for Cisgender and Transgender Communicators". Specific questions were: 1) What are the acoustic and nonverbal communication-based predictors of masculinity-femininity ratings; and 2) Is there a difference in masculinity-femininity ratings between audio only, visual only, and audiovisual (AV) presentation modes? The first question was answered using standard multiple linear regression. Masculinity-femininity ratings collected in the AV presentation mode served as the criterion variable and the full complement of communication variables served as predictor variables. Differences in ratings between stimulus presentation modes were identified using a one-way repeated measures analysis

of variance (ANOVA). Mean femininity rating served as the dependent variable (DV) and presentation mode served as the independent variable (IV). Study 2 (Chapter 3) is under review with the *Journal of Speech, Language, and Hearing Research*.

Chapter 4 describes relationships between communication variables and ratings made by the transgender women participants regarding the femininity of their communication, their satisfaction with their communication, and their QoL. Specific questions included: 1) How are acoustic measures related to communicator-rated a) voice femininity?, b) voice satisfaction?, c) VRQoL?; 2) How are gestural measures related to communicator-rated gestural femininity?; 3) How are acoustic and gestural communication measures related to communicator-rated a) communication femininity?, b) communication satisfaction; and c) overall QoL? Pearson correlation was used for the two analyses involving QoL scores (i.e., the total scores from the TVQ^{MtF} and the QLI), whereas Spearman correlation was used for the remaining analyses involving femininity and communication satisfaction ratings. Study 3 (Chapter 4) will be submitted for publication to the *American Journal of Speech-Language Pathology*.

Finally, gender differences in upper body gestures were identified in the fourth study, which is presented in Chapter 5. Specific questions related to this aim were: 1) Are there differences in upper body nonverbal communication behaviors (i.e., hand movements, hand position, elbow position) between groups based on a. Attributed gender? b. Gender positioning?; and 2) Are there differences in head movements between groups based on a. Attributed gender? b. Gender positioning? Group differences were identified using a

series of one-way ANOVAs with measures of upper body movements serving as the respective DVs and gender groups serving as the IVs. Study 4 (Chapter 5) will be submitted for publication to *Sex Roles: A Journal of Research*.

Methods

Methods common to all of the studies are presented here to provide context for the individual papers. All four studies drew from the same large data set and data collection occurred across two phases. In phase I, communication and QoL data were collected from a group of communicators. Observer ratings were collected from a group of raters in phase II.

Participants

Two groups of participants were recruited to take part in the research: communicators and raters. The series of studies received ethical approval from the University of Alberta Health Research Ethics Board (HREB) (Pro00060133) and operational approvals from the Northern Alberta Clinical Trials and Research Centre (NACTRC) and the Covenant Health Research Centre prior to participant recruitment. All participants provided written consent to participate in the research and were compensated $15 to cover travel expenses.

Communicators. A total of 42 communicators were recruited for the study: 22 transgender women as well as 10 (each) cisgender men and women. Cisgender participants were included to ensure adequate variability in the sample. The mean age of the communicators was 40.93 years (SD = 14.48). Inclusion criteria required that all communicators be fluent English speakers and not have a neurogenic communication disorder. Additional inclusion criteria for transgender women communicators required that they identify as transsexual or transgender and be living in their affirmed gender role

(i.e., woman) the majority of the time (i.e., at least 80% of the time) for at least six months. These criteria are similar to those used in Pasricha, Dacakis, and Oates (2008) and were used to ensure that participants had adequate experience in the affirmed gender role. One transgender participant was excluded from the study due to the presence of a non-North American accent: we believed the accent would impact subsequent acoustic measures and may have acted as a covariate during rating. Another transgender participant was excluded as a result of failing to meet the inclusion criterion of living in the affirmed gender role the majority of the time. These exclusions resulted in a final sample size of n = 40.

The proposed sample size is the same as that used by Wolfe, Ratusnik, Smith, and Northrop (1990) and is representative of the number of transgender women that could reasonably be expected to be recruited from the Edmonton area[3]. Sample size calculations based on planned data analyses suggested recruiting between 81 and over 1700 communicators, depending on the size of the entered minimal difference between variables. This number was not feasible given the small population of transgender women.

[3] Edmonton is a city in Western Canada of just over one million people (Statistics Canada, 2016). The majority of residents are non-immigrants [largely of British, Eastern European (e.g., Ukrainian, Polish), or Canadian First Nations ancestry, given Canada's former immigration policies]. Approximately one quarter of residents are immigrants to Canada from (in descending order) Asia, Europe, Africa, the Americas, and other places of birth.

All communicators were recruited from the city of Edmonton and surrounding areas using convenience sampling, poster advertisements, and word of mouth. Transgender women were recruited from gender and voice clinics, support groups, gender and sexual minority organizations, and from the general community via a media release.

Acoustic and movement data were collected from the communicator group. QoL, communication satisfaction and self-rated communication femininity were collected from the transgender communicators only.

Raters. A total of 20 raters were recruited for participation in the research from the University of Alberta and Edmonton area and represented the communication partners encountered in the communicators' daily lives. Raters were required to (a) be proficient in English; (b) pass a pure-tone hearing screening at 25 dB HL for 500, 1000, 2000, and 4000 Hz bilaterally; (c) have no uncorrected concerns with vision per self-report; (d) have no specific training in listening; (e) have no more than incidental experience listening to or communicating with persons having communication disorders; and (f) have no identified language, learning, or cognitive disabilities per self-report. The mean age of the raters was 29.26 (SD = 9.14) years.

Phase I Data Collection: Communication Data

The first phase of data collection occurred over an 11-month period. Accrual commenced on June 3, 2016 and was completed on May 16, 2017. Phase I data were collected at the Syncrude Centre for Motion and Balance at the Glenrose Rehabilitation Hospital in Edmonton. Upon arrival, communicators provided consent to participate and demographic information. Next they watched a short Pink Panther Cartoon, "In the Pink of the Night" (Davis, 1969). Then they were fitted with recording equipment and

recorded retelling the cartoon to one of the researchers (i.e., Ms. Hardy). Transgender participants completed subjective ratings and the QoL measurements either prior to watching the cartoon or following the cartoon description. More detailed descriptions of the procedures are provided in the sections that follow.

Recording equipment and procedures. Communication data were collected via simultaneous acoustic and motion capture recordings. Acoustic data were obtained via a Shure Mx185 condenser microphone that was attached to the communicator's forehead at a fixed distance of 10 cm from the mouth and connected to an audio-buddy amplifier and 32-bit HP laptop computer. Monaural audio recordings were collected using TF32 software (Milenkovic, 2001) at a sampling rate of 44.1 KHz. Gain was calibrated on the audio-buddy amplifier prior to recording the research tasks. No further adjustments were made once those tasks commenced. Calibration tones also were recorded as in Fox and Boliek (2012), for later use in calculating dBSPL.

Motion data were captured using a Motion Analysis Corporation 8 camera optical motion capture system. A custom set of 15 instrumented motion analysis markers were attached to the head, sternum, seventh cervical vertebra, shoulders, elbows, wrists, and hands [Figure 1 in Study 4 (Chapter 5)]. Once fitted with recording equipment, the communicators were seated on a raised stool facing and at fixed distance from Ms. Hardy and recordings commenced.

Cartoon description task. Recordings were made of the communicator's voices and movements while they recounted the story of a short cartoon. Cartoons are recommended stimuli for eliciting gestures during narrative tasks (D. McNeill, 1992) and the cartoon selected for the present line of research has been used in studies comparing the use of

gestures in monolingual and bilingual speakers. In addition to promoting the use of gestures, a story retell task also controls for use of potentially-confounding vocabulary (e.g., my wife) that may be used inadvertently in a traditional narrative task, while still allowing for more conversational speech patterns than may be observed during a reading task. Tasks such as a picture description, which potentially could meet the same criteria in terms of language used, likely would not result in the required use of gestures, with the exception of pointing to specific parts of the picture. Nicoladis et al. (2007) coded a total of 37 scenes in the six-minute cartoon that potentially could be retold. These authors found that bilingual speakers typically included 15-20 of these scenes when re-telling the story in their first language.

Additional voice recordings were obtained of the communicators reading the carrier phrases'Say hVd again' (with the vowels /i/, /ɑ/, and /u/) and sustaining the vowel /ɑ/ (as in 'paw') for use in specific acoustic analyses. The carrier phrases were the same as those used in Gelfer and Bennett (2013) and were selected to ensure there were common phrases spoken for all participants that contained the vowel /i/ ('ee' as in 'meet').

Acoustic measures. Acoustic analyses were completed using a customized PRAAT script (Kelley, 2016, 2017) or Multi-Speech™ acoustic analysis software. Specific acoustic measures included: mean f_o, mean minimum frequency (Hz); mean maximum frequency (Hz); mean RoS (syllables/s); mean SPL (dBSPL); mean ST range, percent upward intonation shifts; mean average formant frequency (i.e., an average of the first, second, and third vowel formants) for the vowel /i/ (Hz); mean NHR; mean shimmer (%); and mean SPI. All acoustic variables were measured in the context of connected speech (i.e., the cartoon description or a carrier phrase) with the exception of the voice

stability/periodicity measures, which were measured in the context of the sustained vowel (/ɑ/). A 30-45 second section of each communicator's cartoon description was isolated for use in the acoustic analyses. This section terminated approximately one minute prior to the end of the description at the nearest utterance boundary. More detailed descriptions of the acoustic measures are presented in Chapter 2. "Acoustic Predictors of Gender Attribution, Masculinity-Femininity, and Vocal Naturalness Ratings Amongst Transgender and Cisgender Speakers".

Measures of movement. Movement variables included: total head movement in degrees along the mid-sagittal plane (MS Head); total head movement in degrees along the coronal plane (COR Head); total head movement in degrees along the axial plane (AXI Head); total movement in millimeters (mm) of the lead hand irrespective of direction (Hand Mvt); percent of time the lead hand was in the palm-up versus palm-down position (Palmup%); and mean percent elbow: shoulder width (Elbow%). These measures were chosen to represent head nodding, head tilting, head shaking, and hand gesturing behaviors, as well as hand and elbow position, respectively. These variables were measured from the motion capture data obtained during the cartoon description using a customized MATLAB script. The same 30-45 second section of the description was used to measure the motion variables as was used to measure the acoustic variables. More detailed descriptions of the movement variables and motion measurement procedures are presented in Study 2 (Chapter 3) "Contributions of Voice and Nonverbal Communication to Perceived Masculinity-Femininity for Cisgender and Transgender Communicators" and Study 4 (Chapter 5) "Let Me See Your Body Talk: An Exploration of Gender Differences in Upper Body Nonverbal Communication Behaviors".

Measurement of QoL, femininity, and communication satisfaction. Transgender women communicators completed the generic version of the *Quality of Life Index© (QLI)* (Ferrans & Powers, 1998), the *Transsexual Voice Questionnaire (Male-to-Female)© (TVQ^{MtF})* (Dacakis & Davies, 2012), and a series of simple Likert scale ratings of the perceived femininity of their own voice, gestures, and overall communication as well as their satisfaction with their voice and overall communication. Examples of these measures are included in Appendices A-G, respectively. The *QLI* was used to measure satisfaction with aspects of life participation and the perceived importance of these aspects. An overall (total) *QLI* score was obtained and used as the DV representing QOL. The TVQ^{MtF} was used to measure each individual's opinion of their voice and its impact on their life. The total score from obtained from this tool represented VRQoL.

Phase II Data Collection: Rating Data

Upon arrival, raters consented to participate in the research, had their hearing screened, and provided demographic information. They subsequently were seated in front of a laptop computer and provided with a set of noise-cancelling headphones. Once seated, they engaged in a training protocol to become familiar with examples of different voices and gestures (e.g., very feminine, very masculine, ambiguous, natural, unnatural), the presentation modes, and rating procedures. Research trials commenced following the completion of training. Raters were blind to the inclusion of transgender communicators and were debriefed following the completion of the research tasks.

Rating task stimuli. The 30-45 second audio recordings of the communicators performing the cartoon retell task served as the audio only (A) stimuli. Point-light displays, created from the motion data (i.e., .trc files) using a customized software

program called Gender Finder ("Gender Finder," 2017), served as the visual only (V) stimuli (see Figure 1.1). The audiovisual (AV) stimuli contained both the audio tracks and the PLDs.

Figure 1.1. Point-Light Display

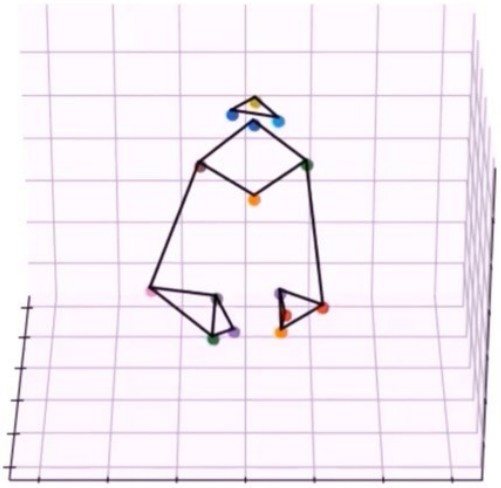


Figure 1.1. An example of a PLD as it appeared to raters.

Using PLDs for the visual stimuli controlled for any potential influence of surface or tissue differences between the communicators with the added benefit of maintaining their anonymity. Structural cues contained in differences in size between speakers was controlled by normalizing the PLDs so they are all the same size while maintaining their original proportions and angles between joints. Gender Finder completed this standardization during the creation of the PLDs. Gender Finder also was used to present the research stimuli to the raters and collect their ratings.

Training protocol. Each participant engaged in a standardized training protocol presented on a laptop computer. They were introduced to relevant research concepts, provided with examples of voices and gestures, and familiarized with the rating procedures via a PowerPoint presentation. Example voices and gestures were obtained

from the previously-rated samples used in Hardy et al. (2016) as well as from individuals identified by the research team to represent the different levels of the criterion measures. After hearing and seeing the example stimuli, the raters completed practice trials using the Gender Finder program. The procedure for the practice trials was identical to the research trials except it used fewer and different stimuli. Raters were permitted to repeat the practice trials as many times as needed in order to feel comfortable with the task. Once raters felt comfortable, they commenced the research trials.

Rating task. Gender Finder presented the research stimuli to raters using a block randomization procedure. The order of the A, V, and AV blocks as well as the individual stimuli within each block were randomly determined for each rater, thus controlling for order effects. Ten percent of the samples in each block were repeated for reliability purposes and remained constant for all raters. A total of 44 stimuli were presented and rated in each block. Gender Finder presented the rating window on the laptop screen after each stimulus item finished playing. Gender attribution ratings were made by selecting from a drop-down list of three choices (i.e., man, woman, can't decide) whereas masculinity-femininity and naturalness ratings were made using direct magnitude estimation scale without modulus (DME-WM). An example of the rating window is presented in Figure 1.2. Raters were asked to rate gender attribution and masculinity-femininity in all presentation modes but only rated vocal naturalness in the audio only mode. Responses were made directly on the computer and stored in an excel file for later analysis.

Figure 1.2. Rating Screen for Audio Only Stimuli

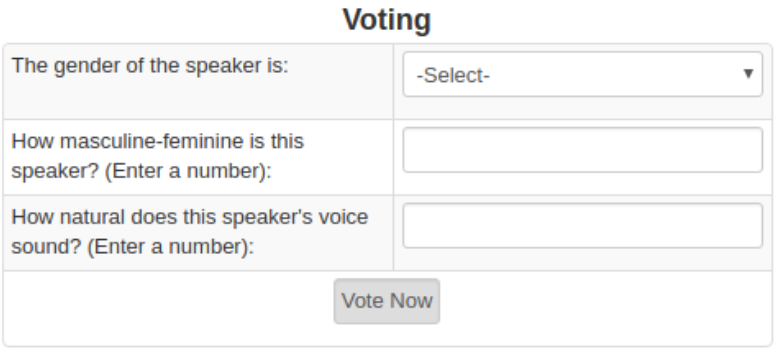


Figure 1.2. An example of the rating window that was presented to raters after each stimulus item in the audio only mode. Voting boxes in visual only and audiovisual modes did not include masculinity-femininity rating fields.

Direct Magnitude Estimation-Without Modulus scales require raters to assign numbers to the samples they are rating in comparison to the first sample played in the series. Raters assign any number they wish to the first sample using their own internal set of critieria and each subsequent sample is given a value relative to that rating. For example, if raters perceived a given sample to be half as natural as the first sample (assigned a value of '100' for ease of explanation), they would assign that sample a value of '50'. Similarly, if it was perceived to be twice as natural it would be assigned a value of '200'. Direct magnitude estimation scales do not have fixed end points and are considered to provide ratio-level data (Kreiman, Gerratt, Kempster, Erman, & Berke, 1993; Whitehill, Lee, & Chun, 2002); however, it should be noted that mathematical conversions of data must first be performed (i.e., modulus equalization) with DME-WM in order to remove the variability secondary to the use of different moduli (Snow & Williges, 1998).

In choosing rating scales, it has been suggested that the nature of the parameter of interest first be determined as this dictates the type of measurement scale that may be used. For example, metathetic measures are perceptual continua that are perceived to vary

in quality (e.g., pitch) and can be divided into equal intervals easily and reliably by raters. For that reason, both Equal Appearing Interval (EAI) and DME scales can be used to rate metathetic variables. In constrast, prothetic continua are perceived to be additive in nature, to vary in amount or quantity (e.g., loudness). Raters are reportedly not able to perceptually divide prothetic measures into equal intervals, making the use of EAI scales inappropriate (Stevens, 1975). Since many communication parameters have been found to be prothetic continua, it is prudent to use a DME scale in order to mitigate the risks and associated bias of using the wrong scale.

The modulus equalization procedure described in Snow and Williges (1998) was used to convert DME-WM ratings rating into ratio-level data and to "account for variance attributable to listeners' individual choice of modulus" (Whitehill et al., 2002, p.84) prior to completing statical analyses. First, each rater's modulus is calculated by converting their original ratings into the logarithmic value and then obtaining the mean. Next, the common modulus (i.e., the mean of all adjusted ratings) is calculated and subtracted from each individual modulus. A normally distributed data set with a common modulus is then created by adjusting each score by the difference between the individual modulus and the common modulus for each respective rater. Finally, the antilog of each item in the new data set is taken to create a ratio-level data set.

Impact of the Dissertation

Previous research has focused predominantly on acoustic measures of the voice, especially f_o. This dissertation research investigated a unique constellation of acoustic variables and other communication factors, such as upper body gesture, that have not yet been quantified in transgender women communicators. Moreover, it advanced our

understanding of the complex relationships between communication features and both observer- and communicator-rated outcome measures such as gender attribution, QoL, and communication satisfaction. The results of this research will add to the evidence base informing voice and communication modification training for TG/GD people.

CHAPTER 2: ACOUSTIC PREDICTORS OF GENDER ATTRIBUTION, MASCULINITY-FEMININITY, AND VOCAL NATURALNESS RATINGS AMONGST TRANSGENDER AND CISGENDER SPEAKERS

Hardy, T. L. D., Rieger, J. M., Wells, K., & Boliek, C. A. (2018). Acoustic predictors of gender attribution, masculinity−femininity, and vocal naturalness ratings amongst transgender and cisgender speakers. *Journal of Voice*, 1–16. Advance online publication. http://doi.org/10.1016/j.jvoice.2018.10.002

Abstract

Purpose: This study aimed to identify the most salient set of acoustic predictors of 1) gender attribution; 2) perceived masculinity-femininity; and 3) perceived vocal naturalness amongst a group of transgender and cisgender speakers to inform voice and communication feminization training programs. This study used a unique set of acoustic variables and included a third, androgynous, choice for gender attribution ratings.

Method: Data were collected across two phases and involved two separate groups of participants: communicators and raters. In the first phase, audio recordings were captured of communicators (n = 40) during cartoon retell, sustained vowel, and carrier phrase tasks. Acoustic measures were obtained from these recordings. In the second phase, raters (n = 20) provided ratings of gender attribution, perceived masculinity-femininity, and vocal naturalness based on a sample of the cartoon description recording.

Results: Results of a multinomial logistic regression analysis identified mean fundamental frequency (f_o) as the sole acoustic measure that changed the odds of being attributed as a woman or ambiguous in gender rather than as a man. Multiple linear regression analyses identified mean f_o, average formant frequency of /i/, and mean sound

pressure level (SPL) as predictors of masculinity-femininity ratings and mean f_o, average formant frequency, and rate of speech (RoS) as predictors of vocal naturalness ratings. **Conclusion:** The results of this study support the continued targeting of f_o and vocal tract resonance in voice and communication feminization/masculinization training programs and provide preliminary evidence for more emphasis being placed on vocal intensity and RoS. Modification of these voice parameters may help clients to achieve a natural-sounding voice that satisfactorily represents their affirmed gender.

Introduction

The discrepancy between gender identity (i.e., internal concept of one's being a man or a woman) (Stryker, 2008) and gender attribution (i.e., how one's gender is perceived by others) (Stryker, 2008) experienced by many transgender women can have negative consequences on quality of life (QoL) (Hancock, 2017), life participation (Bauer & Scheim, 2015; Byrne, 2007; Pasricha et al., 2008), and even personal safety (Bauer & Scheim, 2015; Grant et al., 2011; James et al., 2016b). Consequently, developing a gender expression that is congruent with gender identity can be an important part of the gender role transition process and often motivates transgender women to seek voice and communication feminization training from speech-language pathologists (SLPs). The SLP helps these individuals modify aspects of their communication that influence how others perceive them (E. Coleman et al., 2011). Appropriate selection of training targets requires that decisions be informed by research evidence.

Sound of the voice is known to be a salient cue to gender attribution (Holmberg et al., 2010; Leung et al., 2018; Skuk & Schweinberger, 2014). More specifically, speaking fundamental frequency (f_o) and vocal tract resonance (as measured by vowel formant

frequencies) appear to be particularly important cues; this holds true for both transgender and cisgender speakers (Dacakis, Oates, & Douglas, 2012; Gelfer & Bennett, 2013; Hillenbrand & Clark, 2009; Leung et al., 2018; Oates & Dacakis, 2015). Several other aspects of communication have been proposed as cues to gender attribution such as measures related to intonation, rate of speech (RoS), loudness, and voice quality (Adler, Hirsch, & Mordaunt, 2006; Dacakis et al., 2012; Gelfer & Bennett, 2013; Hancock, Colton, & Douglas, 2014; Holmberg et al., 2010; Owen & Hancock, 2010). The relative contribution of each of these variables to gender attribution currently is not well understood, especially when considering transgender women speakers. Understanding the relative contributions of voice and speech variables to gender attribution in cisgender and transgender women speakers is the focus of the present study.

Speaking Fundamental Frequency

It has been well established in both transgender and cisgender literature that f_o plays an important, perhaps even critical, role in gender attribution (Dacakis et al., 2012; Gelfer & Mikos, 2005; Gelfer & Tice, 2013; Hillenbrand & Clark, 2009; Holmberg et al., 2010; E. J. M. McNeill, Wilson, Clark, & Deakin, 2008; Oates & Dacakis, 1997; Van Borsel, De Cuypere, & Van den Berghe, 2001; Wolfe, Ratusnik, Smith, & Northrop, 1990). In fact, results of a meta-analysis exploring the contribution of f_o to gender attribution showed that this measure accounted for 41.6% of the variance in attributed gender (Leung et al., 2018). This voice measure also has been positively related to femininity ratings made by both unfamiliar listeners and transgender women, themselves (Carew et al., 2007; Gelfer & Tice, 2013; Gorham-Rowan & Morris, 2006; Hardy et al., 2016; Owen & Hancock, 2010). It is important to note, however, that although a good deal of

the variance in gender attribution measures was explained by f_o in the meta-analysis, there remained 58.4% to be explained by other communication factors. Furthermore, there is evidence to suggest that the role of these other factors may be more pronounced during connected speech tasks (Gelfer & Bennett, 2013) and when f_o falls within a 'gender ambiguous' or neutral range (Gelfer & Bennett, 2013; Gelfer & Tice, 2013; Van Borsel et al., 2001), defined as 150-185 Hz (Mordaunt, 2006). This same range has been suggested as a gender-acceptable target for communication feminization training (Gelfer & Mordaunt, 2012). As a result, it is important to identify these other factors and to explore their interaction with f_o in contributing to measures of gender attribution.

Vocal Tract Resonance & Vowel Formant Frequencies

Vowel formant frequencies have long been suspected to play a key role in distinguishing voices perceived as belonging to men and women (Gelfer & Bennett, 2013; Gelfer & Tice, 2013; Deborah Günzburger, 1995; Hillenbrand & Clark, 2009; Mount & Salmon, 1988) and now are generally accepted to be contributors to gender attribution (Dacakis et al., 2012). They also have been related to vocal naturalness ratings for transgender women (Hardy et al., 2016). What remains unclear, however, is the salience of each vowel formant and the strength of the overall contribution.

Studies have shown relationships between auditory-perceptual ratings of gender and/or femininity and each vowel formant (F1, F2, F3) in isolation (Carew et al., 2007; Gelfer & Bennett, 2013; Gelfer & Tice, 2013; Deborah Günzburger, 1995; Mount & Salmon, 1988) as well as between those ratings and an average of the three formants together (Gelfer & Bennett, 2013). However, significant relationships have only been identified between these variables when measured in the context of connected speech and not

isolated vowel productions (Gelfer & Mikos, 2005; Gelfer & Schofield, 2000; Hardy et al., 2016; King et al., 2012). Moreover, vowel formant frequencies account for more of the variance in gender attribution ratings when measured in connected speech than in isolated vowels (Gelfer & Bennett, 2013; Gelfer & Mikos, 2005). Studies investigating relationships between auditory-perceptual ratings and vowel formant frequencies should, therefore, use connected speech as the capture task.

Despite the compelling evidence for the importance of vowel formant frequencies in contributing to gender attribution and potentially natural-sounding voices, they do not seem to be sufficient for consistent correct gender attribution even when combined with f_o (Hillenbrand & Clark, 2009). For that reason, it is important that other acoustic measures be explored as potential contributors as well.

Intonation

Intonation or melody is another aspect of communication that commonly is believed to differ between masculine and feminine communication styles and contribute to natural-sounding speech. Research has shown that transgender women who were attributed as women used more upward intonation shifts, greater frequency variability and semitone (ST) range, and had a higher upper limit for f_o in connected speech tasks such as reading, picture description, and responding to questions (Gelfer & Schofield, 2000; Hancock et al., 2014; Van Borsel et al., 2001). Higher upper limit for f_o also was found to significantly positively correlate with femininity ratings (Gelfer & Schofield, 2000). These studies suggest that more feminine communication styles are associated with greater variability in intonation. Other studies, however, have either found the opposite

relationship (i.e., femininity ratings were significantly related to a narrow ST range) (Owen & Hancock, 2010) or failed to find relationships between intonation measures and femininity ratings at all (Hancock et al., 2014).

Taken together, these studies suggest that the contribution of intonation measures to gender attribution is not yet fully understood. Such measures may have some influence on gender attribution, particularly parameters such as ST range and percentage of upward intonation shifts. Further research is needed to confirm previous results and determine the relative salience of intonation as a cue to gender attribution. Research also is needed in other areas related to prosody such as speaking rate.

Rate of Speech

The cisgender literature is inconclusive with respect to gender differences in RoS. Some studies suggest that men's speech rate is faster than women's in both reading and conversational contexts (Fitzsimons et al., 2001; Verhoeven et al., 2004), whereas others report RoS to be similar between the genders (Tsao & Weismer, 1997; Tsao, Weismer, & Iqbal, 2006; Van Borsel & De Maesschalck, 2008).

There are relatively few data with respect to RoS in transgender women speakers and the data that do exist reflect the same inconsistencies seen in the cisgender literature. For example, early studies found transgender women speakers used a slower RoS when reading words and sentences in the feminine versus masculine gender presentation (Günzburger, 1989, 1995), suggesting they believed a slower RoS sounded more feminine. Conversely, Van Borsel and Maesschalck (2008) found no differences in RoS between cisgender men, cisgender women, and transgender women speakers when reading a standardized passage.

Of note, none of the studies involving transgender women investigated differences in RoS as a function of gender attribution or masculinity-femininity, which are more relevant to communicative success than gender status. Moreover, previous studies used reading tasks and did not investigate the role of RoS in the more ecologically valid conversational context. Further research is needed to explore RoS as a potential gender marker.

Sound Pressure Level

According to Gelfer and Young (1997), sound pressure level (SPL) (perceptual correlate = loudness) is one of the most salient perceptual features of the voice and has been shown to differ in small but significant ways between genders in cisgender speakers. For example, in a conversational context, men speak two to three dBSPL louder than women (Boonin, 2012; Gelfer & Young, 1997) and have a wider conversational loudness range, which relates to use of greater sound pressures in their inflections (Gelfer & Young, 1997).

Although this parameter has a presence in clinical discussions of gender-related communication differences to be considered in voice and communication feminization training (Adler et al., 2006; Dacakis et al., 2012; Oates & Dacakis, 1983, 1997), very little research has investigated the role of SPL in gender attribution, especially as it relates to transgender speakers. The few studies that have been published have found louder voices to be associated with masculine gender presentations (Günzburger, 1989, 1995, 1993) and gender attributions (Holmberg et al., 2010) in connected speech tasks. These results were based on a small number of participants (n = 4 to 6).

Voice Quality & Vocal Perturbation

Voice quality is another, widely studied, communication construct that has been proposed to serve as a cue for gender attribution. There are mutliple ways to perceptually characterize voice quality; however, breathiness and roughness seem to be of particular interest in the context of gender attribution. Feminine-sounding voices have been characterized by increased breathiness (Andrews & Schmidt, 1997; King et al., 2012; Klatt & Klatt, 1990; Mendoza et al., 1996; Södersten et al., 1995; Van Borsel et al., 2009), whereas rougher or 'croakier' voices (i.e., self-rated vocal fry) have been associated with masculine-sounding voices (Holmberg et al., 2010).

Several measures have been used in previous research to quantify these subjective constructs. For example, measures of cycle-to-cycle variability in vocal fold timing (i.e., jitter, relative average perturbation [RAP]) and amplitude (i.e., shimmer) have been related to perceived vocal harshness or hoarseness (King et al., 2012; Wendahl, 1966a, 1966b), Noise-Harmonic Ratio (NHR) with roughness (Bhuta et al., 2004), and Soft Phonation Index (SPI) with breathiness (Bhuta et al., 2004).

There are limited studies investigating these acoustic parameters as a function of gender attribution amongst transgender women. Of note, shimmer previously has been moderately negatively correlated ($p = .076$) with listener femininity ratings (Owen & Hancock, 2010) and SPI has been shown to contribute to an individual being attributed as a woman (Porter, 2012). In other cases, vocal perturbation measures (i.e., jitter, shimmer, NHR) either were excluded from the study or found to correlate weakly with femininity ratings due to a lack of variability in these acoustic measures (Hardy et al., 2016; King et al., 2012; Owen & Hancock, 2010).

In addition to its potential relationship with femininity ratings, shimmer also has been negatively related to naturalness ratings in the pre-training voices of transgender women (n = 25) (Hardy et al., 2016). Although not found to be a significant predictor of naturalness ratings on its own, shimmer did significantly predict these ratings when included in a model with F2 of /a/ and minimum frequency in a maximum phonational frequency range task.

Overall, it appears that voice quality may contribute to perception of gender, femininity-masculinity, and/or vocal naturalness, but additional research is needed to more clearly describe this contribution. Given previous research results and auditory-perceptual characterizations of masculine and feminine voices, shimmer, NHR, and SPI may be the acoustic measures most worthwhile to investigate.

In summary, the existing literature provides compelling evidence for the role of f_o and vocal tract resonance in gender attribution; however, it also is clear that other aspects of the voice serve as cues for gender and masculinity-femininity to others. Research has been somewhat equivocal with respect to the role and relative strength of other acoustic variables as predictors of these auditory-perceptual constructs.

Studies have looked at certain of these acoustic variables either in isolation or in groupings related to a few areas of voice (e.g., f_o, intontation, resonance). A more comprehensive investigation into the relative contribution of acoustic variables - representing a number of different aspects of voice and speech - to gender attribution is needed. Furthermore, although it is becoming increasingly common, not all studies have included perceptual ratings. When included, these ratings often have focused on either gender attribution or perceived masculinity-femininity but not always both. Information

also is lacking describing how acoustic variables contribute to a natural-sounding voice: only one study has included this auditory-perceptual measure with transgender communicators (Hardy et al., 2016).

Purpose

The purpose of this study is to identify the most salient set of acoustic predictors of gender attribution, perceived femininity, and perceived vocal naturalness in a conversational context. Findings will help inform voice and communication feminization training so that goals are centered on the strongest cues to gender in conjunction with the areas identified as most important to the individual client. Improving our understanding of the predictive relationships between these acoustic variables and perception of femininity and vocal naturalness will guide further refinement of gender presentation so that speakers are attributed not only as their affirmed gender but also in a way that is congruent with their specific identity (e.g., a very feminine woman, somewhat masculine woman) without sounding unnatural. In this way, clinicians will be better able to tailor training so that it is appropriate to the self-identified goals of the client.

Specific questions related to these research purposes are: What acoustic measures predict: (1) gender attribution (as man, woman, or ambiguous in gender)?; (2) perceived masculinity-femininity?; and (3) perceived vocal naturalness?

Methods

Participants

Two groups of participants were recruited to take part in the study: communicators and raters. This study received ethical approval from the University of Alberta Health Research Ethics Board (HREB) and operational approvals from the Northern Alberta Clinical Trials and Research Centre (NACTRC) and the Covenant Health Research Centre prior to participant recruitment.

Communicators. A total of 42 communicators were recruited for the study: 22 transgender women as well as 10 (each) cisgender men and women. Inclusion criteria required that all communicators be fluent English speakers and not have a neurogenic communication disorder. Additional inclusion criteria for transgender women communicators required that they identify as transsexual or transgender and be living in their affirmed gender role (i.e., woman) the majority of the time (i.e., at least 80% of the time) for at least six months. One transgender participant was excluded from the study due to the presence of a non-North American accent: the authors believed the accent would impact subsequent acoustic measures and may have acted as a covariate during the rating. Another transgender participant was excluded as a result of failing to meet the inclusion criterion of living in the affirmed gender role the majority of the time. These exclusions resulted in a final sample size of n = 40.

The sample size (for transgender women) used in the present study is the same as that used by Wolfe, Ratusnik, Smith, and Northrop (1990) and is representative of the number of transgender women that could reasonably be expected to be recruited from the Edmonton area. Cisgender men and women were included to ensure adequate variability

in the sample and each of these participants was age-matched with one of the transgender participants. All but two age matches were within two years and the remaining two matches were within four years. There were no significant differences in age between gender identity groups as determined by a one-way between-subjects Analysis of Variance (ANOVA) $F(2,37) = .010$, $p = .990$. Demographic information for all communicators can be found in Table 2.1. Transition-related information for transgender communicators is summarized in Table 2.2.

Table 2.1 *Demographic Information of Communicators*

Demographic variable	Transgender women (n = 20)	Cisgender women (n = 10)	Cisgender men (n = 10)	All communicators (n = 40)
Mean age in years (SD)	41.20 (14.38)	40.90 (14.76)	40.40 (15.93)	40.93 (14.48)
Age range in years	16-69	24-68	18-70	16-70
History of smoking (%)	45	40	30	40
Current smoker (%)	15	0	20	12.5

Table 2.2 *Transition-Related Information for Transgender Communicators*

Transition-related characteristic	N	Mean (SD)	Min-max	*n*
Time living in feminine gender role at least part-time (years)	20	9.08 (11.71)	1.0–52.0	-
Time living "full-time" in feminine gender role (years)	20	7.63 (11.72)	0.75–52.0	-
Length of SLP services (months)	19	1.46 (3.42)	0.0–12.0	-
Time since SLP services (months)	8	45.00 (54.81)	0.0–156.0	-
Receiving HRT	20	-	-	18
Underwent GAS	20	-	-	8
TVQ^{MtF} Total Score	20	60.70 (18.96)	31-91	-

Note. One participant received SLP services but could not recall for how long. Twelve participants had not received SLP services at the time of participation. HRT = hormone replacement therapy. GAS = Gender affirmation surgery. TVQ^{MtF} = Transsexual Voice Questionnaire (Male-to-Female) [53]. Minimum and maximum possible for total score = 30-120.

Information about smoking history was collected due the potential of smoking to affect voice quality (Martins, Tavares, & Pessin, 2017). History of smoking was similar across gender groups and all groups had members who had succeeded in smoking cessation by the time of data collection. Of note, no cisgender women continued to smoke whereas there were similar, small numbers of active smokers in the other two groups. Overall, relatively few of the communicators were smokers.

The TVQ^{MtF} is a self-report tool developed specifically for transgender women as a means of measuring the impact of their voices on their day-to day-lives (Dacakis et al., 2017a). The average total score for this group of transgender women fell 15 points below the midpoint, indicating that, in general, these participants had some negative experiences related to voice function and/or social participation but these difficulties were not necessarily happening on a very frequent basis.

Cisgender communicators were recruited using poster advertisements and word-of-mouth. Transgender communicators were recruited from gender and voice clinics, support groups, gender and sexual minority organizations, and the general public using poster advertisements, word-of-mouth, and a media release. Letters of invitation to participate also were sent to individuals who had been referred for voice and communication feminization services through the voice program at the local rehabilitation hospital. Individuals who were interested in participating and who were awaiting or were actively receiving services were asked to contact another author (CB) in order to prevent feelings of coercion as the first author was the clinician responsible for providing said services. Individuals who already had received services or who were discharged without being seen were asked to contact the first author if they were interested in participating.

Raters. A total of 20 raters were recruited for participation in the study from the University of Alberta and Edmonton area through convenience sampling and word-of-mouth and represented the naïve communication partners encountered in the communicators' daily lives. Half of the raters identified as women and half identified as men. The mean age (n =19) was 29.26 and ranged from 18 – 46 years. One rater declined

to provide her age. Raters were required to (a) be proficient in English; (b) pass a pure-tone hearing screening at 25 dB HL for 500, 1000, 2000, and 4000 Hz bilaterally; (c) have no uncorrected concerns with vision per self-report; (d) have no specific training in listening; (e) have no more than incidental experience listening to or communicating with persons having communication disorders; and (f) have no identified language, learning, or cognitive disabilities per self-report.

Most of the raters had at least some undergraduate education and all had at least a high school diploma. The level of education attained (including partial completion) ranged from community college to advanced graduate level. Raters varied with respect to the amount of their family, friends and/or close colleagues who identified with the greater lesbian, gay, bisexual, transgender, and queer (LGBTQ) community; however, the majority had at least a few. Raters were asked to indicate their sexual orientation on a nine-point scale anchored with 1 = Very Heterosexual and 9 = Very Homosexual as in Hancock and Pool (Hancock & Pool, 2017). The median sexual orientation for the whole group was straight according to the criteria used by those authors (i.e., less than three was considered straight); however, half of the raters scored themselves as a three or greater, which would be considered non-straight by the same criteria. Raters were asked to provide information about their exposure to the LGBTQ community and their sexual orientation because femininity ratings made of transfeminine speakers have been related to sexual orientation (Hancock & Pool, 2017) and because exposure to sexual and gender diverse communities may result in broader conceptualizations of gender expression.

Data Collection for Communication Variables

Recording equipment and set-up. Data collection took place at the Syncrude Centre for Motion and Balance at the Glenrose Rehabilitation Hospital in Edmonton. Upon arrival, participants provided consent to participate in the study and subsequently were fitted with recording equipment. A Shure Mx185 condenser microphone was attached to the communicator's forehead at a fixed distance of 10 cm from the mouth and connected to an audio-buddy amplifier and 32-bit HP laptop computer. Monaural audio recordings were collected using TF32 software (Milenkovic, 2001) at a sampling rate of 44.1 KHz. Gain was calibrated on the audio-buddy amplifier prior to recording the research tasks. No further adjustments were made once those tasks commenced. Communicators were seated on a raised stool and were positioned facing a listener (the first author) as if in a conversational exchange.

Speaking tasks. Audio recordings were taken of the communicators while they retold the story of the Pink Panther cartoon, "In the Pink of the Night" (Davis, 1969). A story retell task was chosen to control for the use of potentially-confounding vocabulary (e.g., my wife) that may inadvertently be used in a traditional narrative task while still allowing for more conversational speech patterns than may be elicited during reading.

Two additional speaking tasks were performed and recorded for use in obtaining specific acoustic measures. First, communicators were asked to read the carrier phrase, 'Say hVd again' (with the vowels /i/, /ɑ/, and /u/), as in Gelfer and Bennett (2013). They read the carrier phrases five times in total. This task was chosen to control phonetic context during vowel production in connected speech. Second, communicators were recorded while sustaining the vowel /ɑ/ (as in 'pop') for approximately 5 seconds over

five trials. This task was included for use in obtaining vocal stability measures, which provide information about non-volitional cycle-to-cycle variability in frequency and SPL (Baken & Orlikoff, 2000). As such, sustained vowels provide an ideal context for capturing vocal stability measures because they do not include volitional variations in frequency and SPL found in connected speech secondary to intonation (Baken & Orlikoff, 2000).

Calibration tones were recorded at the end of each data collection session, as in Fox and Boliek (Fox & Boliek, 2012), for later use in calculating SPL. Calibration tones were generated at 440.0 Hz using a tone generator application on a smartphone. The speaker volume on the smartphone was set to maximum and the speaker was placed at the center of the communicator's mouth (i.e., the same point from which microphone distance was measured).

Editing cartoon descriptions. Communicators were variable in the length and completeness of their cartoon descriptions. Consequently, the narratives were edited to control for length and linguistic complexity. Linguistic complexity was measured in t-units, where a t-unit was defined as "one main clause plus whatever subordinate clauses happen to be attached to or embedded within it" (Hunt, 1965, p.305). Each sample was edited to be between 30-45 seconds in length and contain 8-12 t-units. Care was taken to exclude vegetative acts (e.g., laughing) from the sample. These edited descriptions were used for the rating phase of data collection and for the majority of the acoustic measures.

Acoustic measures. A total of 13 acoustic measures initially were captured: mean f_o (Hz); mean minimum frequency (Hz); mean maximum frequency (Hz); mean RoS (syllables/s); mean SPL (dBSPL); mean ST range, percent upward intonation shifts; mean

first, second, and third vowel formants for the vowel /i/ (Hz); mean NHR; mean shimmer (%); and mean SPI. These served as the predictor (i.e., independent) variables in the statistical analyses. Customized PRAAT (Boersma & Weenink, 2017) scripts (Kelley, 2016, 2017) were used to obtain all the acoustic measures with the exception of ST range and SPI, which were obtained using the Real-Time Pitch and Multi-Dimensional Voice Program modules in Multi-Speech^{TM} (KayPentax, 2008), respectively. The first seven measures were taken from the edited cartoon description, averaged across t-units. Vowel formant frequencies were measured in the context of the carrier phrase at the midpoint of the vowel (averaged across the five productions), and the three vocal stability measures were obtained from the middle one second of the sustained /ɑ/ (averaged across the five productions).

Intonation shifts were defined (Gelfer & Schofield, 2000) as "a change in frequency, with or without interruption of phonation, of at least two semitones" (p.26). For each t-unit, the script calculated the difference between the maximum and minimum frequency and converted the value to semitones using the formula: $30.86314 * \log10(f_o max/f_o min)$ as in Hancock, Colton, and Douglas (2014). If this value was at least two ST, it was counted as an intonation shift. The direction of the shift was determined by calculating the difference between the time of the maximum and minimum f_o values. If the value was positive, it was counted as an upward shift and if negative, it was counted as a downward shift. Percent upward intonation shifts subsequently were calculated by dividing the number of upward shifts by the total number of shifts and multiplying by 100.

All acoustic measures were repeated for eight (20%) randomly selected participants for reliability purposes. Intra-measurer reliability was determined using repeat measures made by the first author whereas inter-measurer reliability was calculated using repeat measures made by the first author and two masters-level SLP students.

Data Collection for Auditory-Perceptual Ratings

Rating procedure and rating scales. Data collection took place in a quiet room either at the Clinical Sciences Building at the University of Alberta or at the Glenrose Rehabilitation Hospital. After providing consent and demographic information, the raters were provided with a standard training protocol via a slideshow presentation. The training introduced raters to relevant concepts (e.g., masculinity-femininity, vocal naturalness) and the rating scales, and provided them with examples of different voices (e.g., voice attributed as man/woman/ambiguous in gender).

Data collection was accomplished using a customized software program, "Gender Finder" ("Gender Finder," 2017) that randomly presented the edited cartoon description samples from all 40 communicators followed by 10% of the files repeated for reliability purposes. Raters were asked to make three ratings for each audio file: gender, perceived masculinity-femininity of the communicator, and perceived vocal naturalness. Gender attribution ratings were made by selecting one of three options from a drop-down menu: man, woman, or can't decide. The third option 'can't decide' was included to expand the binary choice typically provided when inquiring about gender and to capture those presentations that are attributed gender-neutrally since that is a desired outcome for some transgender people. These options were believed to adequately represent gender attributions made by the general public. Masculinity-femininity ratings provided

additional information about the spectrum of gender presentations and how they are perceived. Masculinity-femininity and naturalness ratings both were made using direct magnitude estimation scales without modulus (DME-WM) where higher numbers represented more feminine and more natural voices and lower numbers represented more masculine and more unnatural voices, respectively. This type of scale requires raters to assign a number to the first stimulus item to which they are presented and rate subsequent stimuli in comparison to the first. For example, if the first item was assigned 200 and the second item was perceived to be twice as feminine, it would be assigned 400. Conversely, if the second file was perceived to be half as feminine, it would be assigned 100. The naturalness scale was structured in an analogous fashion. Raters were permitted to use any number of their choosing as long as it was not a negative number or zero. Ratings were entered directly into Gender Finder via a rating window that appeared on the screen following each audio file.

DME scales do not have fixed end points and are considered to provide ratio-level data (Kreiman et al., 1993; Whitehill et al., 2002). Mathematical conversions were performed (i.e., modulus equalization) with the DME-WM ratings prior to their use in statistical analyses to remove the variability secondary to the use of different moduli between raters (Snow & Williges, 1998). That is, ratings were mathematically converted to be on the same scale with the same modulus. Modulus equalization also creates a normal distribution, allowing for the use of parametric statistics.

Raters were given an opportunity to practice using the DME-WM rating scale during the training exercise by assigning numbers to lines of various lengths similar to Snow and Williges (1998). They were provided with additional experience prior to commencing the

research trials during a practice session with the Gender Finder program. The software was written to have a practice module as well as an experimental module. The practice module was the same as the experimental module except that it used a smaller set of unique stimuli and rating data were not stored. The recordings used for the practice module were of two pilot participants from the research team and the two communicators who did not meet inclusion criteria. The practice module could be repeated as many times as the rater desired until they were comfortable with the program and the tasks. No rater completed the practice module more than once.

Raters were seated at a desk or table directly in front of a laptop computer that was loaded with the Gender Finder software and provided with a set of Bose QuietComfort 35 noise-cancelling headphones. All training, practice, and research trials were completed on the laptop computer. Raters were instructed to adjust the volume on the laptop to a comfortable level during the practice trials and were not allowed to make further adjustments once the research trials commenced. They also were provided opportunities to ask questions after the training and practice activities. Research trials began following the practice module. The auditory-perceptual ratings took approximately 35-50 minutes to complete.

Once all ratings were completed, raters were asked to answer a two-item post-rating questionnaire about their exposure to the LGBTQ community and their sexual orientation. They then were debriefed about the study and the reason for asking the post-rating questions. Raters were not advised of the inclusion of transgender women communicators prior to their participation in the study.

Auditory-perceptual ratings. Composite gender attribution, masculinity-femininity, and vocal naturalness ratings were calculated for each communicator. Gender attribution was assigned according to the criteria used in Gelfer and Bennett (2013) where an individual was considered to be consistently attributed as a particular gender if they were attributed that gender at least 80% of the time. If none of the three gender choices was selected at least 80% of the time, that communicator was assigned 'can't decide' as their gender was ambiguous to the raters. Masculinity-femininity and vocal naturalness ratings were averaged across raters to arrive at a mean overall rating. These ratings served as the criterion (i.e., dependent) variables in the statistical analyses.

Statistical Analyses

The first research question regarding acoustic predictors of gender attribution was answered using logistic regression due to the categorical nature of the criterion variable. Standard multiple linear regressions were used to identify sets of acoustic predictors of perceived masculinity-femininity and vocal naturalness ratings and, thus, answer the second and third research questions. Multiple regressions were used because masculinity-femininity and naturalness ratings were continuous variables providing ratio-level data.

Given the large number of predictors and the relatively small sample size, it was necessary to decrease the number of predictors used for each regression analysis to maximize the participant: predictor ratio. This culling of predictors was accomplished by identifying and removing highly correlated predictors and by using a purposeful selection of covariates model-building strategy as recommended by a consulting statistician and by Hosmer, Lemeshow, and Sturdivant (2013). Purposeful selection identifies the most parsimonious set of predictors for use in further analyses. Variables that were highly

correlated with other variables were removed from the potential predictor set for all of the analyses. Purposeful selection of covariates was performed separately for each of the regression analyses.

Results

Reliability of Acoustic Measures

Intra- and inter-measurer reliability were examined for all 13 acoustic measures. Intraclass Correlation Coefficient (ICC) estimates and their 95% confidence intervals (CI) were calculated using SPSS statistical software version 24 based on a two-way mixed effects model with absolute agreement. Results are presented in Table 2.3. Based on the 95% CI and using criteria suggested by Koo and Li (2016), intra-measurer reliability was excellent for all measures except ST range and percent upward intonation shifts. Reliability ranged from good to excellent for ST range and percent upward intonations shifts were not reliable. The reason for lack of reliability was thought to be due to the method of measurement. That variable was excluded from subsequent analyses, including inter-measurer reliability.

Inter-measurer reliability was excellent for seven of the remaining 12 measures. Mean minimum frequency and F1 /i/ ranged from good to excellent reliability whereas F2 /i/, ST range, and mean SPL ranged from moderate to excellent reliability.

Table 2.3 *Intraclass Correlation Coefficients (ICC) and 95% Confidence Intervals (CI) for Acoustic Measures*

	Intra-measurer reliability		Inter-measurer reliability	
Acoustic Measure	ICC	95% CI	ICC	95% CI
f_o	1.0	[.999, 1.0]	.998	[.995, 1.0]
Minimum frequency	.999	[.994, 1.0]	.963	[.887, .992]
Maximum frequency	1.0	[.998, 1.0]	.998	[.994, 1.0]
F1 /i/	1.0	[.998, 1.0]	.949	[.844, .989]
F2 /i/	.983	[.925, .997]	.929	[.792, .984]
F3 /i/	.998	[.987, 1.0]	.989	[.964, .998]
ST range	.978	[.891, .996]	.91	[.741, .979]
% Upward intonation shifts	.699	[(-).562, .940]	-	-
SPL	.999	[.994, 1.0]	.858	[.617, .966]
Shimmer	.996	[.979, .999]	.984	[.947, .996]
NHR	.999	[.993, 1.0]	.979	[.935, .995]
SPI	.997	[.984, .999]	.985	[.953, .997]
RoS	.995	[.975, .999]	.997	[.992, .999]

Reliability of Auditory-Perceptual Ratings

Intra-rater reliability was calculated between the first and second ratings of the four repeated files using percent agreement for gender attribution and two-way mixed effects model ICC estimates with absolute agreement for masculinity-femininity and vocal naturalness ratings. Results indicated that intra-rater reliability was excellent for gender attribution (% agreement = 92.5%), excellent for masculinity-femininity ratings (ICC = .946, 95% CI = .916-.965), and ranged from good to excellent for naturalness ratings (ICC = .875, 95% CI = .804-.921).

Inter-rater reliability was calculated between the ratings made by each of the 20 raters using Fleiss' Kappa for gender attribution and two-way mixed effects model ICC estimates with absolute agreement for masculinity-femininity and vocal naturalness ratings. Results indicated that inter-rater reliability was good for gender attribution (κ = .81), excellent for masculinity-femininity ratings (ICC = .94, 95% CI = .909-.964) and ranged from good to excellent for naturalness ratings (ICC = .854, 95% CI = .777-.913).

Correlations Between Acoustic Variables

Bivariate Pearson Correlation Coefficients were calculated between each of the 12 remaining acoustic variables. Statistically significant correlations (r = .887-.969, $p < .01$) were revealed between f_o, mean minimum frequency, and mean maximum frequency. Given the wealth of literature supporting f_o as a cue to gender attribution, it was selected to remain in the set and the other two variables were excluded.

No other correlation coefficients exceeded 0.8; however, significant correlations were identified between F1 and F2 of /i/ (r = .339, $p < .05$) and F2 and F3 of /i/ (r = .713, p <.01). In an effort to further reduce the number of potential predictors, an average was taken of the three vowel formants of /i/, similar to Gelfer and Bennett (2013), resulting in the creation of a single measure to represent vocal tract resonance (i.e., "average formant frequency"). Descriptive statistics for each of the remaining eight acoustic measures along with the DME-WM scale ratings are presented in Table 2.4. Of note, SPI data were missing for five of the communicators (12.5%) because all five trials of the sustained vowel yielded an error response when analyzed in Multi-Speech (KayPentax, 2008). In addition, one rater (R12) did not rate naturalness for any communicator. Two additional

naturalness ratings (from R2 and R7) were missing and likely were due to mouse click errors during rating. These omissions represented 5.25% of the total naturalness rating data.

Correlations Between Transition-Related Factors and Auditory-Perceptual Ratings

Bivariate Pearson Correlation Coefficients were calculated between a number of personal and transition-related factors and masculinity-femininity ratings for the transgender communicators. There were no significant relationships between the ratings and time since transition commenced (r = .225, p = .341, two-tailed), length of time living fulltime in the feminine gender role (r = .212, p = .370, two-tailed), length of voice and communication training (r = -.207, p = .394, two-tailed), time since the completion of training (r = 674, p = .067, two-tailed), or TVQ^{MtF} total score (r = -.381, p = .097, two-tailed).

Table 2.4 *Descriptive Statistics for Acoustic Measures, Femininity, and Naturalness Ratings*

	All Communicators (n = 40)		Attributed Man (n = 25)		Attributed Woman (n = 11)		Attributed Ambiguously (n = 4)	
Acoustic Measure	Mean (SD)	Min-Max	Mean (SD)	Min-Max	Mean (SD)	Min-Max	Mean (SD)	Min-Max
f_o (Hz)	147.73 (40.04)	96.22-230.58	122.93 (21.67)	96.22-196.28	194.60 (22.14)	150.46-230.58	173.84 (36.84)	146.06-227.93
Average formant frequency (Hz)	1952.22 (141.04)	1692.99-2344.64	1887.16 (95.55)	1692.99-2116.39	2101.73 (138.82)	1875.80-2344.64	1947.66 (63.83)	1864.57-2000.87
Shimmer (%)	0.035 (0.017)	0.014-0.077	0.041 (0.018)	0.014-0.077	0.037 (0.012)	0.014-0.058	0.024 (0.004)	0.018-0.028
NHR	0.017 (0.12)	0.002-0.137	0.021 (0.03)	0.002-0.137	0.012 (0.01)	0.003-0.034	0.007 (0.01)	0.002-0.011
SPI[a]	18.39 (8.84)	4.73-34.36	16.95 (7.80)	4.73-34.06	19.13 (11.54)	7.09-34.36	24.64 (6.10)	16.20-29.88
SPL (dBSPL)	74.68 (3.50)	69.39-82.43	75.02 (3.57)	69.76-82.43	73.77 (3.86)	69.39-82.19	75.13 (1.88)	73.46-76.85
RoS (syllables/s)	3.70 (0.56)	2.36-4.81	3.69 (0.52)	2.60-4.81	3.94 (0.56)	2.81-4.53	3.15 (0.56)	2.46-3.68
ST range	9.0 (2.59)	4.58-16.63	9.21 (2.65)	5.17-16.63	8.23 (2.71)	4.58-14.25	9.79 (1.72)	7.63-11.67
Masculinity-femininity rating	21.51 (7.48)	10.72-35.71	-	-	-	-	-	-

Vocal naturalness rating	17.64 (1.98)	11.12-19.94	-	-	-	-	-	-

Note. [a] SPI had missing values and resultant sample sizes were: all communicators = 35; attributed man = 22; attributed woman = 9; attributed ambiguously = 4.

Acoustic Predictors of Gender Attribution

A multinomial logistic regression was used to model the relationship between a set of acoustic predictors and gender attribution (man, woman, can't decide) (n = 40). Purposeful selection univariable analyses (i.e., 1-way ANOVA or Kruskal-Wallis test) identified four significant variables ($\alpha = .25$): f_o, average formant frequency, shimmer, and RoS. These variables were retained for inclusion in the regression analysis, resulting in a sample size of $n = 10$ per variable.

A total of 25 (62.5%) communicators were attributed as men, 11 (27.5%) were attributed as women and raters could not decide the gender of four (10%) of the communicators; therefore, the referent group for the analysis was 'attributed as a man' because it was the most frequently attributed gender. The Goodness of Fit test was not significant χ^2 (70, $N = 40$) = 47.437, $p = .995$, indicating that the model fit the data well. The final model (i.e., the model including the predictors) predicted gender attribution significantly better than the model that contained only the intercept χ^2 (8, $N = 40$) = 47.437, $p < .001$. Pseudo R^2 (Nagelkerke) revealed that 83.9% of the variance in gender attribution was explained by the model. As is shown in Table 2.5, f_o and RoS were statistically significant; however, only f_o had a significant parameter for comparing attributed man and attributed woman (Wald $\chi^2 = 3.951$, df = 1, $p = .047$). The odds that an individual would be attributed as a woman rather than a man increased by a factor of 1.133 (95% CI = 1.002-1.281) for every unit increase in mean f_o. Overall the model correctly predicted gender attribution for 90.0% of cases.

Table 2.5 *Unique Contributions of Predictors in the Multinomial Logistic Regression (n = 40)*

	Likelihood ratio test	
Predictor	χ^2 (df = 2)	p (α = .05)
F_o	15.80	<.01*
Average formant frequency	1.59	.451
Shimmer	3.67	.159
RoS	8.42	.015*

Note. χ^2 = the increase in -2 log likelihood when the predictor is removed from the full model. Statistically significant results are marked with *.

The analysis subsequently was repeated including only f_o as a predictor. Similar results were obtained except that in this case, the model explained 71.2% of the variance in gender attribution and significant parameters were revealed for comparisons between attributed man and both attributed woman (Wald χ^2 = 9.236, df = 1, p = .002) and can't decide (Wald χ^2 = 5.792, df = 1, p = .016). For every unit increase in mean f_o, the odds that an individual would be attributed as a woman or ambiguous in gender increased by a factor of 1.107 (95% CI = 1.037–1.182) and 1.079 (95% CI = 1.014–1.149), respectively. Mean f_o was 122.93 Hz (SD = 21.67) for communicators who were attributed as men, 194.60 Hz (SD = 22.14) for those attributed as women, and 173.84 Hz (SD = 36.84) for those whose gender was ambiguous to raters (see Figure 2.1). This model correctly predicted gender attribution in 85% of cases.

Figure 2.1. Differences in F_o by Attributed Gender

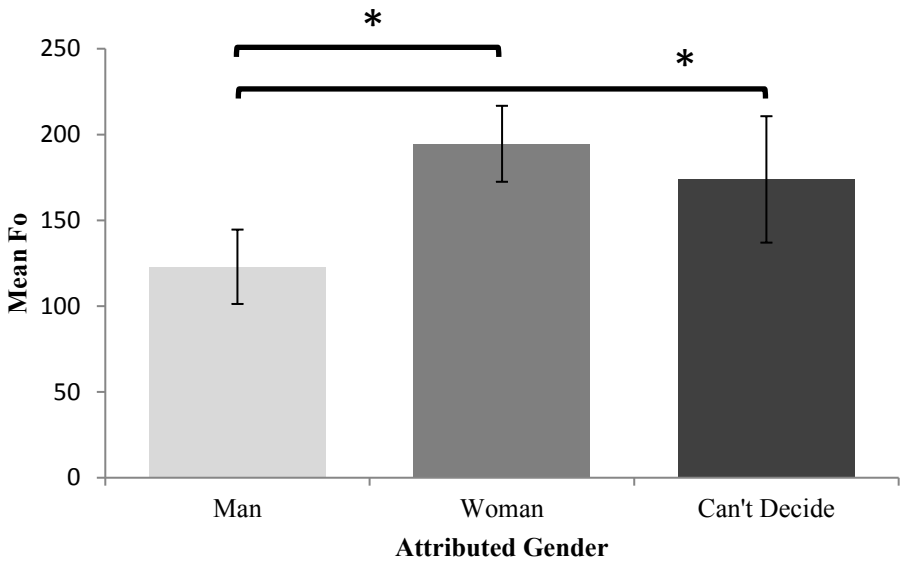


Figure 2.1. Mean f_o for each attributed gender group. Error bars represent +/- 1 SD. Statistically significant group comparisons are marked with *.

Acoustic Predictors of Masculinity-Femininity Ratings

Multiple linear regression was used to identify a set of acoustic predictors of masculinity-femininity ratings. Five predictors were found to be significant through purposeful selection procedures (i.e., simple regression) and thus were retained for the multiple regression analysis using the standard method. These variables included f_o, average formant frequency, shimmer, NHR, and SPL.

The results revealed a significant model: $F(5,34) = 30.154$, $p < .001$ that explained 78.9% of the variance in masculinity-femininity ratings (adjusted $R^2 = .789$). Regression coefficients for the predictor variables are summarized in Table 2.6. F_o, average formant frequency, and SPL were significant predictors of masculinity-femininity ratings. F_o and average formant frequency were positively related to the ratings whereas SPL was

negatively related: Higher speaking frequencies and vocal tract resonance characteristics were associated with more feminine/less masculine ratings and voices with greater SPL were associated with less feminine/more masculine ratings.

Table 2.6 *Unstandardized and Standardized Regression Coefficients for the Predictors Entered into the Model*

Predictor	*B*	SE B	β	*p* (α = .05)
F_o	0.101[a]	0.019[a]	.539[a]	< .001*[a]
	0.104[b]	0.019[b]	.557[b]	< .001*[b]
Average formant frequency	0.019[a]	0.005[a]	.366[a]	.001*[a]
	0.021[b]	0.005[b]	.395[b]	< .001*[b]
Shimmer	-83.175[a]	50.692[a]	-.191[a]	.110[a]
	-[b]	-[b]	-[b]	-[b]
NHR	48.563[a]	33.208[a]	.151[a]	.153[a]
	-[b]	-[b]	-[b]	-[b]
SPL	-0.652[a]	0.172[a]	-.305[a]	.001*[a]
	-0.538	0.160[b]	-.252[b]	.002*[b]

Note. Significant predictors of masculinity-femininity ratings are marked with *. [a] Coefficients obtained from the first model including all five predictor variables. [b] Coefficients obtained from the second model including only the three predictors that were significant in the first model.

The multiple regression was repeated using only the three significant predictors in order to obtain new regression coefficients. These also are presented in Table 2.6. Differences in beta coefficients did not exceed 20%; therefore, there likely was not an interaction with a variable that was excluded (Hosmer et al., 2013). These coefficients revealed that f_o made the greatest contribution to the model, followed by average formant frequency, and finally, SPL. For every change of one SD in mean f_o (i.e., 40.04 Hz), we would predict a change of .557 SD in masculinity-femininity ratings or about four points.

Similarly, for every one SD change in average formant frequency and SPL, we would expect a change of .395 SD (i.e., approximately three points) and -.252 SD (i.e., approximately 2 points) in masculinity-femininity ratings, respectively.

Acoustic Predictors of Vocal Naturalness Ratings

The same statistical procedures were used to identify the set of acoustic predictors of vocal naturalness as were used for identifying predictors of masculinity-femininity ratings. Four predictors were selected to enter into the multiple regression model: f_o, average formant frequency, NHR, and RoS. The results of the multiple linear regression revealed a significant model: $F(4,35) = 6.579$, $p < .001$ that explained 36.4% of the variance in vocal naturalness ratings (adjusted $R^2 = .364$). Significant predictors included f_o, average formant frequency, and RoS. F_o was negatively related to vocal naturalness. The other two predictors were positively related. Regression coefficients are presented in Table 2.7.

Table 2.7 *Unstandardized and Standardized Regression Coefficients for the Predictors Entered into the Model*

Predictor	*B*	SE B	β	*p* (α = .05)
F_o	-0.029[a]	0.009 [a]	-0.592 [a]	.002* [a]
	-0.032[b]	0.009 [b]	-.650 [b]	.001* [b]
Average formant frequency	0.009 [a]	0.002 [a]	0.641 [a]	.001* [a]
	0.009 [b]	0.002 [b]	.642 [b]	.001* [b]
NHR	15.462 [a]	11.511 [a]	0.182 [a]	.188 [a]
	- [b]	- [b]	- [b]	- [b]
RoS	1.091 [a]	0.457 [a]	0.312 [a]	.022* [a]
	1.035 [b]	0.460 [b]	.295 [b]	.031* [b]

Note. Significant predictors of vocal naturalness ratings are marked with *. [a] Coefficients obtained from the first model including all four predictor variables. [b] Coefficients obtained from the second model including only the three predictors that were significant in the first model.

Once again, the multiple regression was repeated with only the three significant predictors entered into the model. The revised regression coefficients also are summarized in Table 2.7. Interactions with NHR are unlikely given the small differences in beta coefficients.

Discussion

This study aims to identify the most salient set of acoustic predictors of gender attribution, as well as perceived masculinity-femininity and vocal naturalness for the purpose of contributing to the evidence base informing voice and communication modification training programs for transgender individuals. Each of these aims was addressed with a separate research question. The discussion will focus primarily on significant findings.

Acoustic Predictors of Gender Attribution

The results of the multinomial logistic regression revealed a significant predictive model and identified f_o as the sole acoustic predictor that significantly changed the odds of being attributed as a woman or ambiguous in gender rather than as a man. As mean f_o increased, so too did the chances of being attributed a gender other than man.

Mean f_o for participants attributed as men and women was consistent with reported norms for cisgender speakers (Baken & Orlikoff, 2000). Mean f_o for the group attributed as gender ambiguous was within the suggested gender ambiguous or gender neutral range (Mordaunt, 2006). Examination of the minimum and maximum values (Table 2.4) revealed that mean f_o for participants attributed as women did not fall below 150 Hz. This value is at the lower boundary of the gender-neutral range and is slightly lower than minimum threshold values previously suggested for a transgender woman to be attributed as a woman, approximately 155-160 Hz (Spencer, 1988; Wolfe et al., 1990). The minimum value for mean f_o for the ambiguously attributed group (146.06 Hz) was just below the gender neutral frequency range. As such, targeting a mean f_o that falls within the cisgender woman or gender neutral range appears to be beneficial for minimizing attributions as a man based on the voice.

Maximum mean f_o values also were informative. That of participants attributed as men was approximately 30-35 Hz lower than those of the other two attributed gender groups; however, it fell within the cisgender woman range. This result provides further evidence for the assertion that modifying f_o alone or speaking at a mean f_o within the cisgender

woman range may not be sufficient for consistent attribution as a woman. In this study, 71.2% of the variance in gender attribution could be accounted for by f_o, leaving 28.8%, almost a third of the variance, remaining to be explained by other variables.

It is possible that other acoustic variables may have emerged as significant predictors had the sample had more variability with more even distribution across the three attributed gender groups. Almost 2/3 of the communicators who participated in this study were attributed as men, including 16 of the 20 transgender communicators (one was attributed as a woman and three were attributed as gender ambiguous). Only about 1/4 were attributed as women, and 1/10 were attributed ambiguously. Two recent review articles (one with meta-analysis) provide direction regarding which of the acoustic variables included in the present study may have emerged as significant predictors of gender attribution. These variables include vowel formant frequencies, f_o variability, minimum and maximum f_o, intonation measures such as directional shifts and ST range, SPL, and acoustic correlates of voice quality (e.g., SPI) (Leung et al., 2018; Oates & Dacakis, 2015). Each of these variables has evidence to support its status as a voice marker of speaker gender.

Methodological factors also may have limited the identification of significant predictors in addition to the lack of sufficient variability in communicator voice and speech characteristics. Several of the listed variables were measured in the present study but not included in the regression analyses either because they were measured for descriptive purposes only (i.e., f_o variability), the measurements were not reliable (i.e.,

percent upward intonation shifts), or they were excluded due to multicollinearity with other variables (i.e., minimum and maximum f_o). It would be valuable to systematically explore the salience of these predictors along with f_o amongst gender diverse speakers.

Despite the limitations in communicator variability and acoustic variable inclusion, the results of the present study nevertheless are consistent with previous research in demonstrating that f_o is a strong cue to gender attribution (Leung et al., 2018) and support the common and recommended practice of targeting raising f_o in voice and communication feminization training (Davies, Papp, & Antoni, 2015; Hancock & Garabedian, 2013).

Acoustic Predictors of Masculinity-Femininity Ratings

The results of the multiple linear regression analyses identified f_o, average formant frequency, and SPL as significant predictors of masculinity-femininity ratings. Higher values of f_o and average formant frequency were associated with more feminine ratings, whereas higher SPL values were related to more masculine ratings. Together, these variables accounted for most of the variability in masculinity-femininity ratings (adjusted $R^2 = .789$), with f_o contributing the most to the model, followed by average formant frequency, and finally SPL. These results confirm and extend previous research and provide novel findings.

First, the identification of f_o as a predictor of masculinity-femininity ratings is not surprising given the results presented in the previous section and in the existing literature. The results of this study confirm f_o as an important cue to perceived masculinity-femininity of the voice and further support the goal of raising f_o in training when aiming to achieve a more feminine-sounding voice.

Second, the results also are consistent with previously published research that reported vocal tract resonance to be a salient cue for gender attribution in the context of connected speech (Gelfer & Bennett, 2013; Hillenbrand & Clark, 2009) and extends those findings by revealing the same relationship with ratings of masculinity-femininity and including transgender communicators. In addition, the results of the present study are similar to those reported in an unpublished master's thesis (Dahl, 2018). Both studies found positive relationships between vowel formant frequencies and masculinity-femininity ratings. Dahl found stronger relationships for cisgender participants than for transgender ones; however, small sample size (n = 12 transgender communicators, n = 10 cisgender communicators) and an ordinal rating scale may have limited the results.

It is interesting and important to note that formant frequency emerged as a significant predictor in this study even though it was averaged across the first three formants of /i/. This result suggests that vocal tract resonance, overall, is salient for conveying one's masculinity or femininity to others. Average formant frequency has been linked to gender attribution in the past and was suggested to better represent the perceptual experience of listeners than individual vowel formants (R. Coleman, 1976; Gelfer & Bennett, 2013). The previous studies averaged the first three formants for two or three vowels (i.e., /i/, /ɑ/, /ɛ/, and/or /u/) whereas the present study used only /i/. Since the results were the same when the composite measure was obtained using only one vowel as when two or three vowels were used, perhaps clinicians could use the single vowel composite for assessment and outcome measurement in order to save time.

Taken together, these results suggest that training communication behaviors that decrease the length of the vocal tract and/or size of the resonating chambers (e.g., lips spreading, anterior tongue carriage) and thus, raise the frequency of the vowel formants overall, may result in meaningful training outcomes. Such outcomes were observed in a intervention study that targeted these behaviors (Carew et al., 2007) as well as in two other studies investigating post-intervention changes following a course of voice feminization training that included facilitation of oral resonance patterns (Gelfer & Tice, 2013; Gelfer & Van Dong, 2013). All three studies found that participants had higher formant frequencies following training and sounded more feminine/less masculine to listeners. Similarly, Hirsch (Hirsch, 2017) presented a resonance training approach developed to modify vocal tract resonance to sound higher or lower/deeper, depending on the goals of the transgender client. The approach uses physiologic behaviors such as lip spreading/rounding and tongue carriage, as described previously, as well as articulatory precision, and end-of-utterance mouth posture to achieve desired acoustic outcomes. Although not yet supported by efficacy research, the approach reportedly is backed by promising anecdotal evidence. The successful outcomes achieved via Hirsch's training program offer preliminary support for the practical application of the present study's vocal tract resonance findings.

Although the composite average formant frequency measure used in the present study was found to be a significant predictor of perceptual ratings, it may be useful to look at the specific vowel formants in the future in order to ascertain whether or not certain vowel formants are more salient than others. This information would be useful in guiding training programs in terms of knowing which behaviors might be most effective in

altering vocal masculinity-femininity. Modifying vocal tract resonance already is a common training goal (Hancock & Garabedian, 2013) and experts have recommended that it be targeted in conjunction with f_o (Davies et al., 2015). The results of this study provide additional support for these practices and recommendations.

Finally, this is the first study to identify SPL as a significant predictor of masculinity-femininity ratings for a group that includes transgender communicators. It extends the limited evidence provided in the cisgender literature by exploring SPL as a function of auditory-perceptual ratings rather than group differences based on gender identity. Auditory-perceptual ratings are more meaningful for informing voice feminization training because of the relevance to the oft-identified outcome of gender attribution. These results show that not only do men speak louder than women (Gelfer & Young, 1997) but individuals who are perceived to be more masculine speak louder than those who are perceived to be feminine as well. The results of Holmberg and colleagues (2010) suggested such a relationship existed between SPL and gender attribution for transgender communicators but results were limited to only four participants in that study (n=26).

The identification of SPL as a significant predictor of perceived masculinity-femininity provides preliminary evidence to support more deliberate targeting of vocal loudness in voice feminization training programs as a means of refining gender presentation (e.g., somewhat feminine-sounding versus very feminine-sounding). More research is needed to investigate the effect of changing SPL on training outcomes such as gender attribution or masculinity-femininity ratings and to refine training targets. The softest communicator in this study (who also was rated as most feminine) spoke at 69.39 dBSPL (measured at 10 cm from the mouth), more than one SD below the mean of 74.68

(3.50). The loudest communicator was more than two SD above the mean at 82.43 dBSPL. Based on these results, 70 dBSPL may be an appropriate target for achieving a more feminine-sounding voice.

In summary, the results of this research question support the continued focus on modifying f_o and vowel formant frequencies in training and suggest that modifications in SPL may be worthy of more attention. SPL currently is not a common training target (Hancock & Garabedian, 2013).

Acoustic Predictors of Naturalness Ratings

Multiple linear regression results revealed three significant predictors of vocal naturalness ratings: f_o, average formant frequency, and RoS. The variables contributed to the model to a decreasing degree according to the order listed. F_o was negatively related to vocal naturalness ratings whereas both average vowel formant and RoS were positively related to the perceptual rating.

Given that the raters attributed the majority of the communicators (n = 25) as men, it is logical that lower f_o would be associated with higher vocal naturalness ratings overall as raters would be expecting low-pitched voices for speakers who are men. Conversely, high-pitched voices would not be expected for the majority of speakers and may have sounded unusual.

An analogous argument can be used to explain the positive relationship between vocal naturalness ratings and RoS. The studies that have reported gender differences in RoS found that men spoke faster than women (Fitzsimons et al., 2001; Verhoeven et al., 2004). It follows that communicators who spoke faster in this study generally would sound more natural since most of them were thought to be men.

It also is possible that RoS is related to speech naturalness in general and not associated with a particular group of speakers. For example, studies from the stuttering and Alternative and Augmentative Communication (AAC) literature have shown faster RoS to be associated with higher naturalness ratings (Ratcliff et al., 2002; Van Borsel & Eeckhout, 2008). Unfortunately, this same relationship was not established for fluent speakers (Van Borsel & Eeckhout, 2008). In the present study, the mean RoS was 3.7 syllables/s (SD = 0.56), which is within normal limits for adult speakers during extemporaneous speech (Robb et al., 2004). Eight speakers were between one and two SD slower than the group mean and of those, seven spoke at a rate slower than three syllables/s. This rate falls more than two SD below the mean for American English speakers reported in Robb et al. (2004). Five of the speakers who spoke slower than 3 syllables/s were assigned an average vocal naturalness rating that was below the group mean of 17.64 (1.98). Most of the communicators (n = 25) were rated to be at or above average in terms of naturalness. It stands to reason that extremely slow RoS would sound unusual or unnatural to some listeners and as such, should be avoided by those seeking to develop speech patterns characteristic of a typical speaker. More research is needed to confirm or clarify this relationship for transgender communicators.

The positive relationship between average formant frequency and vocal naturalness ratings is more difficult to interpret than the other two relationships. As average formant frequency increased, so too did the naturalness ratings. This result can be explained on the basis of gender attribution for those who were attributed as women (n = 11) because

one would expect to higher vowel formants for feminine-sounding speakers. This argument does not hold, however, for the majority of the communicators who were attributed as men.

Some clarity may be gleaned from Hardy et al. (2016), a study reporting similar findings. Those authors examined acoustic predictors of vocal naturalness for a group of transgender women (n = 25) who had not yet received voice and communication feminization training and found F2 of /ɑ/ to be one of three variables included in the predictive model. It was hypothesized that the transgender women communicators were perhaps speaking in a way that resulted in raised F2 values (e.g., anterior tongue carriage), which made them feel more feminine and, consequently, more comfortable, which in turn made them sound more natural to listeners. That also may have been the case for the transgender communicators included in the present study, who account for half of the participants. As such, between the two hypotheses, the positive relationship can be explained for most of the study participants.

The results of this research question suggest that f_o and average vowel formant are not only important for modifying the perceived masculinity-femininity of the voice but also are important to ensure the voice sounds natural or authentic. Clinicians also should be cognizant of the influence of RoS on vocal naturalness and discourage clients from speaking at an unusually slow rate.

Relationships Between Personal Factors and Auditory-Perceptual Ratings

Given the broad inclusion criteria for the transgender communicators and the aim of the present research to contribute to the evidence-base informing voice and communication feminization training, the potential impact of transition-related variables

is worthy of brief mention. Whereas one might assume that certain of these factors (e.g., length of voice and communication training) are related to auditory-perceptual ratings, in fact, there were no systematic relationships between the ratings and time since transition commenced, length of time living fulltime in the feminine gender, length of voice and communication training, time since the completion of training, or TVQ^{MtF} total score for the participants in this study. These results are in keeping with previous findings reporting a lack of significant relationships between a variety of personal factors and listener-reported auditory-perceptual ratings of voice for transgender speakers (Holmberg et al., 2010; Owen & Hancock, 2010). Smoking may have impacted voice quality for some of the communicators; however, it was not formally tested. Although length of training was not significantly related to auditory perceptual ratings in the present study, it is not known whether those who had participated in training had achieved and/or maintained their training goals. Given that mean length of training was so short (i.e., approximately six weeks), it is possible and perhaps probable that participants had not generalized skills to spontaneous connected speech. In general, the ratings obtained in the present study appeared to be independent of the known personal factors. It is possible that other, unmeasured characteristics may have systematically varied with the ratings.

In summary, the present study successfully identified acoustic predictors of the three auditory-perceptual ratings of interest. F_o was the sole predictor of gender attribution ratings; f_o, average formant frequency, and SPL significantly predicted of masculinity-femininity ratings; and f_o, average formant frequency, and RoS significantly predicted naturalness ratings.

Limitations and Future Directions

This study was limited by the lack of variability in the transgender participants despite the relatively large sample size. As was described in the section "Acoustic Predictors of Gender Attribution", 16/20 transgender participants were attributed as men. Only one was attributed as a woman and three were not attributed consistently as either gender. These results are not completely unexpected given that more than half of these communicators (n = 12) had not yet participated in voice and communication feminization training.

There also may have been a self-selection bias amongst the individuals who volunteered for the study. A concerted effort was made to reach as many transgender women in the community as possible; however, many who were known to have made noticeable changes to their voices (e.g., raised f_o) or were successfully "stealth" (self-described) did not volunteer for the study. A small number made inquiries about the line of research but ultimately chose not to participate. Perhaps those who struggled more with their voices and the attributions made by others felt more compelled to contribute to the research whereas those who were having less difficulty had moved forward to "Getting on with Life" or post-transition stage described by Byrne (2007) and had less interest in being involved. It is possible that other factors such as restrictions on participation time and location also may have contributed to the recruitment challenges.

Another limitation of the present study was sample size. Although large when compared to previous communication studies involving transgender participants, the sample size was relatively small for the complex statistical tests and number of associated predictor variables used. Results would have been more robust and may have revealed

additional significant predictors with a greater sample size: predictor variable ratio, especially in the case of the logistic regression (Tabachnick & Fidell, 2013). Similarly, attributed gender groups were unequal, which is not ideal for logistic regression (Tabachnick & Fidell, 2013).

In summary, other variables may have emerged as predictors of the auditory-perceptual ratings (especially gender attribution) had there been greater representation of feminine-sounding speakers. Future studies should aim to replicate these results with a larger sample size and more even distribution of transgender participants across the spectrum of gender attribution. As a next step, it would be valuable to test the models created in the present study and any replication studies that follow in terms of their ability to correctly predict auditory-perceptual ratings. If accurate, the associated regression equations potentially could be used as clinical outcome measures. It also would be valuable to explore how these acoustic variables relate to communicator-based outcome measures such as satisfaction with voice.

A final limitation of the present study was the exclusion of certain acoustic variables. As mentioned in the section "Acoustic Predictors of Gender Attribution", several variables that have evidence to support their contribution to auditory-perceptual ratings related to gender attribution were measured but not included in the regression analyses. In particular, the exclusions unfortunately resulted in a lack of representation of intonation amongst the predictor variables. Future researchers should determine reliable methods of measuring intonation variables such as percentage of upward intonation shifts used and investigate their salience along with those predictors already identified as significant.

The present study also excluded variables related to other aspects of communication such as nonverbal communication (e.g., gesture use). These behaviors may be addressed in voice and communication modification training (Adler et al., 2006) yet there is a dearth of research exploring their role in cueing gender attribution for transgender communicators. Future studies should systematically measure nonverbal communication behaviors and determine whether they are related to perceptual ratings associated with training outcomes.

Study Contributions and Conclusions

This study used a unique set of acoustic variables to predict both gender attribution and perceived masculinity-femininity and moved away from the gender binary to include participants whose gender was attributed ambiguously. The inclusion of naturalness ratings provided needed insight into the acoustic factors that contribute to a natural or authentic-sounding voice. Results suggest clinicians should continue to target pitch and vocal tract resonance in their training programs and consider a greater focus on vocal loudness adjustments. By addressing these voice characteristics and avoiding speaking very slowly, transgender clients may be able to achieve a voice that satisfactorily represents their affirmed gender.

CHAPTER 3: CONTRIBUTIONS OF VOICE AND NONVERBAL COMMUNICATION TO PERCEIVED MASCULINITY-FEMININITY FOR CISGENDER AND TRANSGENDER COMMUNICATORS

Abstract

Purpose: The purpose of this study was twofold: 1. Expand upon Hardy, Rieger, Wells, and Boliek (2018) by identifying a set of communication-based predictors (including both acoustic and gestural variables) of masculinity-femininity ratings; and 2. Explore differences in ratings between audio and audiovisual presentation modes for transgender and cisgender communicators.

Method: The voices and gestures of a group of cisgender men and women (n=10 of each) and transgender women (n=20) communicators were recorded while they recounted the story of a cartoon using acoustic and motion capture recording systems. A total of 17 acoustic and gestural variables were measured from these recordings. A group of observers (n=20) rated each communicator's masculinity-femininity based on 30-45 second samples of the cartoon description presented in three modes: audio, visual, and audio visual. Visual and audiovisual stimuli contained point-light displays standardized for size. Ratings were made using a direct magnitude estimation scale without modulus. Communication-based predictors of masculinity-femininity ratings were identified using multiple regression and analysis of variance was used to determine the effect of presentation mode on perceptual ratings.

Results: Fundamental frequency (f_o), average vowel formant, and sound pressure level (SPL) were identified as significant predictors of masculinity-femininity ratings for these communicators. No nonverbal communication variables significantly predicted the subjective ratings when entered into a model with acoustic variables. Communicators were rated significantly more feminine in the audio than the audiovisual mode.
Conclusions: Both study purposes were met. Results support continued emphasis on f_o and vocal tract resonance in voice and communication modification training with transgender individuals and provide evidence for the potential benefit of modifying SPL, especially when a masculine presentation is desired. Gestural behaviors may have limited contribution to perceived masculinity-femininity.

Introduction

Voice and communication training is an important part of the gender affirmation process for many transgender or gender diverse (GD) individuals (i.e., individuals whose gender positioning is different than the one assigned to them at birth). Such training assists the individual to achieve communication behaviors that are congruent (based on the client's views and socially-constructed paradigms) with gender positioning to reduce feelings of dysphoria and potentially reduce instances of misgendering by others. Aspects of gender presentation such as the way a person acts and the way a person sounds are believed to contribute to gender attribution (Byrne, 2007) and thus, training targets may include, for example, voice and nonverbal communication behaviors (e.g., gestures) (Adler, Hirsch, & Mordaunt, 2012; Hancock & Garabedian, 2013).

There is a growing body of research including transgender participants investigating voice-related aspects of communication (i.e., acoustic measures) and their contribution to gender attribution and femininity ratings. Recent review articles have reported the importance of speaking fundamental frequency (f_o) and vocal tract resonance characteristics [represented by vowel formant frequency (VFF)] as acoustic gender markers, but suggest that other aspects of communication such as intonation, voice quality, intensity, and rate of speech require more research in order to more clearly understand their role in gender presentation and attribution (Davies et al., 2015; Oates & Dacakis, 2015).

The transgender research literature suggests that mean f_o is positively related to femininity ratings (Hardy et al., 2016; E. J. M. McNeill et al., 2008; Owen & Hancock, 2010) and gender attribution as a woman (Gelfer & Mikos, 2005; Gelfer & Schofield, 2000; Holmberg et al., 2010; King et al., 2012; Leung et al., 2018). VFFs have been shown to be similarly positively associated with such perceptual ratings in connected speech (Carew, Dacakis, & Oates, 2007; Günzburger, 1993; Leung et al., 2018; Mount & Salmon, 1988).

In terms of intonation or inflection, studies have shown large variations in intonation patterns (characterized by frequent use of upward intonation shifts, wide frequency variability and utterance semitone [ST] range, and/or high upper limit of f_o) to be perceived as feminine (Gelfer & Schofield, 2000; Hancock, Colton, & Douglas, 2014; Wolfe, Ratusnik, Smith, & Northrop, 1990) and monotone patterns to be perceived as

masculine (Wolfe et al., 1990). Results, however, have not been consistent. For example, Owen and Hancock (2010) found a negative relationship between utterance ST range and femininity ratings made by unfamiliar listeners.

A recent systematic review of the voice- and speech-related factors contributing to gender attribution reported that breathy and tense voice qualities and avoidance of glottal fry contributed to perceived femininity (Leung et al., 2018). In contrast, rougher or "croakier" voice qualities have been perceived as more masculine (Holmberg et al., 2010). These associations were based on auditory-perceptual measures and relationships with corresponding acoustic parameters have been equivocal. For example, there is some evidence to suggest associations between perceived femininity or attribution as a woman and Soft Phonation Index (SPI) (Porter, 2012) and shimmer (Owen & Hancock, 2010). These vocal perturbation measures have been used to measure breathiness (Bhuta et al., 2004; Owen & Hancock, 2010) and harshness or hoarseness (King et al., 2012), respectively. Other studies have reported a lack of relationship between auditory-perceptual ratings of femininity and a variety of vocal perturbation and signal-to-noise measures (Hardy et al., 2016; King et al., 2012; Owen & Hancock, 2010). Overall, the available evidence exploring relationships between acoustic measures of voice quality and ratings related to gender attribution remains weak given the mixed results reported in the research literature; however, there has been a call for additional studies to advance understanding of the contribution of voice quality to gender attribution (Leung et al., 2018).

There is a paucity of research investigating the roles of loudness or sound pressure level (SPL) and rate of speech (RoS) in cueing gender amongst transgender communicators. Three published studies provide preliminary evidence to suggest that louder voices or higher SPL may contribute to perceived masculinity (Günzburger, 1995; Hardy, Rieger, Wells, & Boliek, 2018; Holmberg et al., 2010). For example, Hardy and colleagues (2018) found that SPL significantly predicted masculinity-femininity ratings of transgender and cisgender speakers' voices during a cartoon retell task and higher SPL was associated with more masculine ratings than lower SPL. These results, and those of the other two available studies, are consistent with studies from the cisgender (i.e., individuals whose gender positioning is consistent with the gender assigned to them at birth) research literature that report gender differences in SPL with men speaking two to five dB SPL louder than women (Brockmann, Drinnan, Storck, & Carding, 2011; Gelfer & Young, 1997). This difference is perceptually meaningful given that the general consensus for just noticeable difference (JND) for intensity is 3 dB (Koffi, 2018). Further research is needed to better elucidate the contribution of this potentially subtle cue to gender attribution.

Only one study (Hardy et al., 2018) has explored relationships between RoS and auditory-perceptual ratings of gender attribution with transgender communicators and a couple have investigated gender positioning or gender presentation differences (Günzburger, 1995; Van Borsel & De Maesschalck, 2008). Hardy and colleagues (2018) found that RoS significantly predicted gender attribution ratings when considered independently but not when assessed with other acoutic variables (e.g., f_o). RoS did not significantly predict masculinity-femininity ratings in that study. Given the study

limitations (e.g., lack of diversity in attribution ratings for trangender participants), additional research is needed to advance understanding of how RoS contributes to perceptual ratings related to gender attribution.

Studies investigating group differences based on gender presentation and gender positioning have reported transgender women used slower RoS when reading in the feminine versus masculine gender presentation mode (Günzburger, 1995) and no significant differences between cisgender men, cisgender women, and transgender women speakers when reading (Van Borsel & De Maesschalck, 2008). Similar inconsistencies are reflected in the cisgender literature. That is, studies have reported either no gender differences in RoS (Tsao & Weismer, 1997; Tsao, Weismer, & Iqbal, 2006; Van Borsel & De Maesschalck, 2008) or that men speak faster than women (Fitzsimons et al., 2001; Verhoeven et al., 2004) in both extemporaneous speech and reading tasks. Whereas differences based on gender positioning are interesting, those based on perceptual ratings are more meaningful for informing voice and communication modification training. As such, additional studies including transgender communicators are needed to elucidate the contribution of RoS to relevant perceptual ratings.

There remains much to be learned about the salience of acoustic measures as cues to perceived masculinity-femininity. It also is important to begin to gain an understanding of the contribution of nonverbal communication behaviors because, as yet, there are no published studies including transgender participants exploring their contribution to gender attribution and femininity ratings (Davies et al., 2015). Both audio and visual information is available during in-person exchanges; therefore, it is important to understand the contribution of both voice- and motion-based cues when taken together.

Indeed, Van Borsel, De Cuypere, and Van den Berghe (2001) found that there was a significant interaction between the voice and physical appearance. Ratings of "femaleness" differed across three presentation modes (audio only, visual only, and audio visual) and ratings were higher when visual cues were present. It should be noted that this study used video recordings of the head and shoulders that allowed raters to see aesthetic presentation (e.g., hair, make-up) as well as physical features (e.g., facial structure) of the participants and did not consider gestures.

Pasricha, Dacakis, and Oates (2008) qualitatively studied a group of transgender women (n=12) regarding their perceptions of and satisfaction with their communication behaviors. The study participants reported that they felt they used feminine body language naturally and that these behaviors were seen as feminine by others. Unfortunately, nonverbal communication behaviors have not yet been objectively measured with transgender participants. The cisgender research literature, however, can provide some guidance in this regard. For example, women use more vertical head movements (i.e., greater in frequency and amplitude) (Ashenfelter, Boker, Waddell, & Vitanov, 2009; Boker et al., 2011; Helweg-Larsen et al., 2004; Jones et al., 1999), more semantically meaningful gestures as opposed to general movement (Hostetter & Hopkins, 2002; Nicoladis et al., 2007), and keep the arms closer to the trunk when gesturing (Hirsch, 2006) than men. It also is possible that women use different gestures than men (e.g., women use palm-up versus palm-down gestures and men use pointing) (Peterson, 1976; Yang, 2010). The differences just described predominantly were based on studies utilizing coding analyses of videotaped or live observations; the general audience work of linguists, communication and body language experts; and expert clinical experience, all

of which may be susceptible to the influence of stereotypic expectancies and confirmation bias (Hamilton, Sherman, & Ruvolo, 1990; Tskhay & Rule, 2013). Another body of research offers, perhaps, less biased information regarding potential gender differences in nonverbal communication behaviors and their contribution to gender attribution and/or perceived masculinity femininity. That research comes from the field of biomotion.

Biomotion utilizes movement-tracking technology that measures behaviors directly. Sophisticated camera/computer systems track the 3-D movements of passive motion markers attached to major body joints. The motion data subsequently can be used to create avatars or point-light displays (PLD). A PLD appears as a constellation of moving dots, which observers are able to recognize as human (Mather & Murdoch, 1994; Wöllner & Deconinck, 2013). Numerous studies have demonstrated that observers are able to accurately attribute an individual's gender based on PLD information in contexts such as walking or orchestral conducting (Kozlowski & Cutting, 1977; Mather & Murdoch, 1994; Pollick, Kay, Heim, & Stringer, 2005; Troje, 2002; Wöllner & Deconinck, 2013). Moreover, perceptual judgments based on PLDs are similar to those that are made when observers can see the actual figure (Abernethy & Zawi, 2007) yet they are able to control for differences in body morphology (e.g., secondary sex characteristics, face) or appearance (e.g., clothing, hairstyle) (Johnson et al., 2011; Pollick et al., 2005). Such control makes them ideal for the study of nonverbal communication behaviors.

The biomotion literature is somewhat limited in its exploration of nonverbal communication behaviors, especially as they relate to gender attribution and perceived masculinity-femininity, but does offer some guidance in this regard. For example, Hill and Johnston (2001) found both rigid and non-rigid head movements presented via an averaged or gender ambiguous animated head contributed to gender attribution ratings during a joke-telling scenario. Rigid head movements are those that do not involve change in shape such as nodding, tilting, and shaking whereas non-rigid head movements do result in change of shape (e.g., facial expression). This study provides additional support for investigating the role of head movements in cueing ratings related to gender attribution.

Motion information also was found to contribute to femininity ratings made of re-synthesized avatars (head and shoulders in view) manipulated to vary in gender presentation (labeled by those authors as sex) in addition to appearance whereas sex ratings were based on appearance alone (Boker et al., 2011). Those authors did not specify what types of movements contributed to the ratings; however, given the view of the re-synthesized videos, they must have been related to head and/or facial movement.

Overall, there remain notable gaps in the literature concerning the contribution of nonverbal communication to gender attribution and/or perceived masculinity-femininity. First, no available studies investigating such contributions have included transgender participants. Second, nonverbal communication behaviors have not yet been objectively quantified in the transgender population. Cisgender studies suggest movement may contribute to gender attribution and/or masculinity-femininity ratings and identify head

and hand/arm movements as variables of particular interest. Furthermore, biomotion research provides evidence that PLDs offer desirable control over appearance-related aspects of gender presentation that may confound ratings of nonverbal communication.

In summary, the available research literature demonstrates multiple inconsistences regarding relationships between communication variables and perceptual ratings related to gender attribution, and is very limited in some areas (e.g., RoS, nonverbal communication) when considering population-specific studies (i.e., those that include transgender participants). Additional research is needed to more fully explain how communication, including both acoustic and gestural parameters, contributes to the way one's gender presentation is perceived by others. Moreover, a better understanding of these communication parameters and their impact on gender presentation will help inform communication modification training. It also is important to explore the influence of presentation mode and corresponding availability of auditory and visual information to begin to understand whether communication behaviors should be trained differently across social settings. For example, if there is no difference between ratings based solely on the voice versus those based on voice and movement/gestural information, it would seem that nonverbal cues do not contribute to said ratings and may be of limited clinical importance as training targets. In that case, voice-based characteristics would be the focus of training for all social contexts e.g., face-to-face, on the phone). Conversely, if ratings do differ as a function of presentation mode, nonverbal behaviors may be worthwhile training targets for contexts involving face-to-face interactions.

The purpose of the present study was to identify a set of communication-based predictors of perceived masculinity-femininity and to explore how one's gender presentation might be perceived differently across different modalities. The present study expanded upon Hardy, Rieger, Wells, and Boliek (2018) by examining a more comprehensive set of communication variables that included both acoustic and nonverbal measures. Including both types of communication variables allowed for an exploration of the unique contributions of voice and gestural characteristics to perceived masculinity-femininity as well as their combined contributions. The study aims were achieved by answering two questions: 1) What are the acoustic and nonverbal communication-based predictors of masculinity-femininity ratings; and 2) Is there a difference in masculinity-femininity ratings between audio only, visual only, and audiovisual (AV) presentation modes? Communication variables selected for inclusion in the present study were those that 1) Can be targeted in voice and communication modification training, and 2) Have at least some evidence to indicate they may play a role in cueing masculinity-femininity.

Methods

This observational study was one in a series of studies investigating the relationships between communication behaviors and perceptual ratings related to gender attribution. Methods for this study were described in detail in Hardy et al. (2018), and are summarized here and expanded to include a description of the motion measurements (e.g., hand gestures). TRIPOD reporting guidelines were adopted for this study (Collins, Reitsma, Altman, & Moons, n.d.). These guidelines were developed for the transparent reporting of studies involving predictive models and provide direction regarding pertinent information to include in reports (e.g., description of how predictors were handled in the

analysis, predictive model). Ethical and operational approvals were obtained for this study from the University of Alberta Health Research Ethics Board (HREB) (Pro00060133) and from the Northern Alberta Clinical Trials and Research Centre (NACTRC) and Covenant Health Research Centre, respectively, prior to participant recruitment.

Participants

Separate sets of participants were recruited for each phase of data collection: a group of communicators (n = 40) for phase I and a group of raters (n = 20) for phase II. Communicators were transgender women (n = 20), cisgender women (n = 10), and cisgender men (n = 10). Each cisgender participant was age-matched with a transgender participant and the mean age for the group was 40.93 years (SD = 14.48). Transgender participants self-identified as such and had been living in their affirmed gender role 80% of the time or more for at least six months. They were not required to have participated in voice and communication modification training and in fact, the majority (60%) had not. Training was not required because there are other means of modifying communication behaviors and additional limitations to sample size were not desirable. Only transgender women were included in the present study as they represent the vast majority of the clinical population currently seeking voice and communication modification services. Communicators were recruited from the city of Edmonton and surrounding areas via convenience sampling, recruitment letters, and a media release.

Raters were individuals who identified as men and women (n = 10 each) and who had adequate hearing, vision, and cognition to complete the rating tasks. Raters were required to pass a pure-tone hearing screening at 25 dB HL for 500, 1000, 2000, and 4000 Hz

bilaterally. Cognitive and visual functioning were based on self-report. As with the communicators, raters were recruited from the same area using convenience sampling. Demographic information is summarized in Table 3.1. As a group, the raters identified as heterosexual according to the criteria described in Hancock and Pool (2017) [i.e., self-rated sexual orientation of 1-2 (inclusive) = heterosexual and 3-9 (inclusive) = non-heterosexual (p. 5)], but individual ratings represented the full spectrum of sexual orientations from very heterosexual to very homosexual (i.e., 1-9).

Table 3.1 *Rater Demographic Information*

Characteristic	All Raters (n = 20)	Women Raters (n = 10)	Men Raters (n = 10)
Mean age in years (SD)	29.26 (9.14)	25.67 (8.19)	32.5 (9.11)
Highest level of education			
High school			
Community college	0 (0%)	0 (0%)	0 (0%)
Undergraduate	2 (10%)	0 (0%)	2 (20%)
Advanced graduate	12 (60%)	8 (80%)	4 (40%)
	6 (30%)	2 (20%)	4 (40%)
FFCC identify as LGBTQ			
None (0%)	3 (15%)	1 (10%)	2 (20%)
Few (<20%)	10 (50%)	4 (40%)	6 (60%)
Many (20-50%)	6 (30%)	4 (40%)	2 (20%)
Most (>50%)	1 (5%)	1 (10%)	0 (0%)
Median sexual orientation	2.75	3.5	1.0

Note. One woman rater declined to provide her age. FFCC = Friends, family, and close colleagues.

Phase I Data Collection: Communication Data

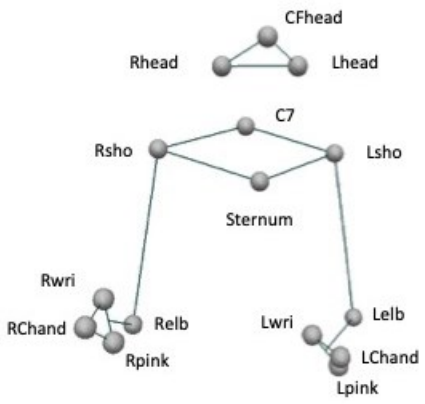
The first phase of data collection occurred over an 11-month period. Accrual commenced on June 3, 2016 and was completed on May 16, 2017.

Recording equipment and procedures. Communication data were collected at the Syncrude Centre for Motion and Balance at the Glenrose Rehabilitation Hospital in Edmonton via simultaneous acoustic and motion capture recordings. Acoustic recordings were obtained at a sampling rate of 44.1 KHz with a Shure Mx185 condenser microphone affixed to the participant's forehead at 10 cm from the middle of the mouth (with lips closed).

Motion data were captured using a Motion Analysis Corporation 8 camera optical motion capture system. The markers placed on the body were 18mm diameter on a 2mm base. Data were collected at a rate of 120 Hz. Any gaps in the data were interpolated using the 'virtual join' function within the Cortex software from Motion Analysis Corporation. This function uses nearby markers to fill gaps and assumes no relative movement is present. Gaps in the tracking data primarily were due to physical occlusion of another body part covering the marker and typically were less than 200ms in duration. Gaps longer than 200ms usually occurred when the hands were placed on the lap. During these phases very little movement is expected and the virtual join function is imperceptible.

The marker set used was a custom set of 15 physical markers and 3 landmarks created using the 'virtual marker' function. The physical markers were placed as depicted in Figure 3.1. Care was taken to ensure that the four wrist markers properly defined the flexion-extension axis of the wrist and that the left and right head markers were level.

Figure 3.1. Physical Marker Set Placement



Marker Name	Description of Location
C7	Spinous process of C7 vertebra
Sternum	Midpoint between the sternoclavicular joints
Lsho, Rsho	Centre of the acromion
Lelb, Relb	Lateral epicondyle of the elbow
Lwri, Rwri, Lpink, Rpink	Radial and ulner styloid process of the wrist
Lhand, Rhand	Midshaft of the third metacarpal
Fhead	Centre of the forehead
Lhead, Rhead	Temporal bone just posterior to the orbit

Figure 3.1. Names and locations of the 15 physical markers.

Participants were seated on a raised stool to prevent leaning on the legs, which would inhibit natural gesturing behaviors. They were positioned facing the first author who was seated on a lower stool and slightly to the communicator's right in order to ensure that the passive motion markers remained visible to the motion capture cameras throughout the recording. Data collection began with the participant looking straight ahead at a mark on the wall that was leveled with the eyes to determine a neutral head position. The CFhead virtual marker was manually adjusted to create a vertical 'face' of the head segment from the neutral position. The participant was asked to hold their arms outstretched laterally in a Tee pose. Wrist angle neutralization was necessary because the marker on the dorsal aspect of the wrist causes the appearance of wrist extension. Neutralization was achieved by using a 30mm depth (from Lhand, Rhand marker joint centers) to determine center hand locations, which were labeled LChand and RChand. After holding the Tee pose for a few seconds, the participant was cued to clap their hands. The clap was used as a

landmark when the audio and motion capture recordings were synchronized. The participant commenced a cartoon description task immediately following the clap. They were instructed to clap again at the end of their description.

Cartoon description and speaking tasks. Communicators watched the Pink Panther cartoon, "In the Pink of the Night" (Davis, 1969) prior to being fitted with recording equipment. Immediately after the recording equipment was in place, participants were asked to retell the story of the cartoon to the first author (a cisgender woman) as if they were telling it to a friend who had not seen the cartoon. Their gestures and voices were recorded while they retold the story. Cartoons are recommended stimuli for eliciting gestures during narrative tasks (D. McNeill, 1992) and this particular cartoon has been used in studies comparing the use of gestures in monolingual and bilingual speakers (Nicoladis, Pika, & Marentette, 2009; Nicoladis et al., 2007; Smithson & Nicoladis, 2014). The communicators voices also were recorded while they said the phrases "say hVd again" (V = /i/, /ɑ/, and /u/) and sustained /ɑ/ for approximately five seconds. These tasks were included for the purpose of measuring vowel formant frequencies and voice stability and periodicity variables, respectively, as per established speech acoustics protocols. Each of the additional speech tasks was repeated five times.

The cartoon descriptions were edited to be of similar length and linguistic complexity across communicators. The resulting samples were 30-45 seconds long and contained 8-12 t-units (the unit of linguistic complexity use in the present study). A t-unit is "one main clause plus whatever subordinate clauses happen to be attached to or embedded within it" (Hunt, 1965, p. 305). The sample terminated at the end of the t-unit that occurred approximately 15 seconds from the end of the description. The beginning of the

sample was determined by locating the point in the description that was 30-45 seconds and 8-12 t-units earlier. These points were chosen to exclude task "warm-up" and the termination of the story. Most participants were describing similar content in their samples. The edited recordings were used for the Phase II rating tasks and for extracting acoustic measurements. No acoustic signal modifications were made to the recordings.

Acoustic measures. A unique set of eleven acoustic measures were captured from the audio recordings using customized PRAAT (Boersma & Weenink, 2017) scripts (Kelley, 2016, 2017) and Multi-Speech^{TM} acoustic analysis software (KayPentax, 2008). These measures included: mean f_o (Hz); mean minimum frequency (Hz); mean maximum frequency (Hz); mean RoS (syllables/s); mean SPL (dBSPL); mean ST range, percent upward intonation shifts; average formant frequency for the vowel /i/ (i.e., an average of the first three formants) (Hz); mean NHR; mean shimmer (%); and mean SPI. All except the last four variables were measured in the context of the cartoon description sample. VFFs were extracted from the midpoint of the /i/ vowel in the carrier phrase and voice perturbation and spectral noise measurements were extracted from the middle one second of the sustained vowel. A mean value was calculated for each variable across all the t-units in the sample, the five productions of the carrier phrases, or the sustained vowels as appropriate. Of note, SPL measurements were adjusted as per Fox and Boliek (2012) using a correction factor calculated from a calibration tone measured during the recording session. Mean minimum and maximum frequencies were found to be highly correlated with mean f_o and subsequently were excluded from the predictor set leaving a total of nine acoustic predictor variables.

Movement measures. Six movement variables were measured from the motion capture recordings: total head movement in degrees along the mid-sagittal plane (MS Head); total head movement in degrees along the coronal plane (COR Head); total head movement in degrees along the axial plane (AXI Head); total movement in millimeters (mm) of the lead hand irrespective of direction (Hand Mvt); percent of time the lead hand was in the palm-up versus palm-down position (Palmup%); and mean percent elbow : shoulder width (Elbow%). These measures were chosen to represent head nodding, head tilting, head shaking, and hand gesturing behaviors, as well as hand and elbow position, respectively. Hand-based measures initially were taken from both left and right sides but only those from the dominant hand (i.e., the hand that moved more) were used for subsequent analyses. MS Head, COR Head, and AXI head ultimately were found to be moderately to highly correlated; therefore, MS Head was selected to represent head movements for this study.

A customized script was created using a commercial software package (MATLAB, 2016) to complete post-processing of the motion capture data. The script extracted information representing each of the six movement variables from the X, Y, and Z coordinate data in the motion capture files (in .trc format). The X coordinate corresponded to the anterior-posterior directions (anterior was negative), Y to left and right (right was positive), and Z to up-down (up was positive). A brief description of each measurement can be found in Table 3.2. It should be noted that head movement measures did not allow for discrimination between frequency and amplitude of movement. That is, the total movement could have been comprised of small and frequent or large and

infrequent movements or some combination of both. Sample time and frame restrictions were added to the script to ensure movements were being measured from the same portion of the cartoon description as the acoustic measures and perceptual ratings.

Table 3.2 *Methods for Measuring Movement Variables*

Movement variable	Mathematical Definition of Movement
MS Head	Total movement of the forehead marker trace calculated as the sum of the angular distances from frame to frame between XZ-coordinates of the forehead marker and a vertically-pointing normal vector.
Hand Mvt	Euclidean distance of the hand marker from frame to frame summed across all frames.
Palmup%	Calculated as the percentage of the time the vector pointing from the palm (cross product of wrist-hand and wrist-pink) pointed up in the world coordinates.
Elbow%	Elbow distance = Euclidean distance between the left and right elbow markers. Shoulder distance = Euclidean distance between the left and right shoulder markers. Calculated elbow distance/shoulder distance * 100 for each frame and then averaged across all frames in the sample.

Script verification.

The script was verified using a multistep approach before and after data analysis to ensure the quality and reliability of the data was high. These steps included, for example, determining agreement of output with known measures and with estimates derived from point-light displays (PLDs). The known measures were obtained by recording the first author performing tasks designed to represent each of the movement variables in the motion lab using the motion capture system described previously. For example, one task involved moving the hand in a straight line for a pre-determined distance (i.e., 60 cm). Script output then was compared to the known value. The first author also estimated values for each of the movement variables based on the motion observed in the PLDs.

These estimates were compared to the script output. Verification was completed using both actual output values and graphic plots of marker movements. All six movement variables were verified using this process.

Phase II Data Collection: Perceptual Rating Data

The second phase of data collection took place over a period of three months. Perceptual ratings were collected between November 23, 2017 and February 18, 2018.

Creating stimuli for rating tasks. Separate stimuli were created for each of the three presentation modes. The audio stimuli consisted of the edited cartoon descriptions described previously. Visual and AV stimuli required the creation of a visual representation of the communicator's movements.

A custom-created software program ("Gender Finder," 2017) was used to create PLDs from the motion capture data files (.c3d) for each communicator similar to that shown in Figure 3.1. The size of the PLDs was standardized to further reduce bias resulting from differences in body morphology. The PLD was merged with the complete audio recording of the cartoon description using Windows Movie Maker. The claps performed by the participants at the beginning and end of the cartoon description were used to ensure the audio and video tracks were synchronized. That is, once the files were merged, they were observed to ensure the sound of the clap corresponded to the point in the PLD when the hands made contact. Once synchronized, the AV files subsequently were edited to represent the same 30-45 seconds used in the audio stimuli. These files served as the AV stimuli. The exact time point of each edit was documented and used to specify the points at which the full PLDs should be edited to create the visual stimuli thus ensuring the same section of the cartoon description was used for all three types of stimuli.

Procedure for perceptual ratings. Data collection took place at the University of Alberta or the Glenrose Rehabilitation Hospital in a room with low ambient noise. Raters first consented to participate and then provided demographic information. Next they were seated at a table or desk in front of a laptop computer and supplied with a set of Bose QuietComfort 35 noise-cancelling headphones. Each rater then engaged in a standardized training protocol delivered by PowerPoint presentation that was designed to familiarize them with research constructs (e.g., masculinity-femininity, naturalness), the rating scale, and the study protocol. The training also provided them with examples of voices (e.g., previously rated as very masculine, very feminine, and somewhere between masculine and feminine) and PLD movements (i.e., made by a cisgender man and a cisgender woman), and an opportunity to practice making ratings. Voice examples were of transgender women and cisgender communicators. The recordings of the transgender women were rated by naïve listeners in Hardy and colleagues (2016). Examples also included recordings of a cisgender woman rated by the research team as having a very feminine voice and of individuals with voice disorders (used in speech-language pathology training) as examples of unnatural voices. The rating practice included a module in Gender Finder that was identical to the research protocol except it used the files of only four unique communicators and did not save the ratings. Once raters were comfortable with the research tasks, data collection with study stimuli commenced.

Rating data were collected using Gender Finder (2017). To control for order effects, the program presented research stimuli using a block randomization procedure. That is, for each rater, the program randomly determined the order of the presentation mode blocks (i.e., audio only, visual only, and AV) and randomly presented the files within

each of the blocks. The files in each block included the 30-45 second cartoon description samples from all 40 communicators as well as 10% of the files repeated for reliability purposes for a total of 44 files in each block. The repeated files were randomly selected (but were the same across blocks and raters) and presented in random order at the end of the block. All raters completed all three blocks. Extensive testing occurred prior to the commencement of data collection to ensure Gender Finder presented files randomly as designed and recorded responses accurately.

Gender Finder (2017) presented the rating platform on the laptop screen following each stimulus item and organized the completed ratings into a spreadsheet. Raters were asked to indicate the perceived masculinity-femininity of the communicator using a direct magnitude estimation scale without modulus (DME-WM) where lower numbers corresponded to more masculine ratings and higher numbers more feminine. Since there is no modulus with DME-WM, raters are instructed to rate the first item as they wish and then rate subsequent items compared to the first. Any number can be used for the rating as long as it is positive (excluding zero). After rating data were collected from all 20 raters, modulus equalization was performed to convert the ratings to a common, normally distributed scale using the same modulus as per Snow and Williges (1998). Modulus equalization involves logarithmic transformations and score-adjustment using a constant derived from the individual and common moduli. The converted ratings were used in the statistical analyses. Each block took approximately 35-50 minutes to complete and raters were permitted to take breaks between blocks as needed.

Finally, raters were asked to provide information about their sexual orientation and ally status (i.e., amount of people in their personal and social circles who identify as a member of the LGBTQ community) because of potential associations with ratings. Hancock and Poole (2017) found that transfeminine speakers received higher femininity ratings from non-heterosexual listeners than heterosexual ones. Although those authors did not find significant differences in ratings based on ally status, it is possible that greater allies also may have experience with a wider range of gender expressions and thus have a broader conceptualization of what behaviors represent masculine or feminine presentation than those who are not allies. Information in the present study was gathered via a two-item questionnaire. Sexual orientation was rated on a nine-point scale anchored with 1 = Very Heterosexual and 9 = Very Homosexual and raters selected ally status from pre-determined percentage ranges representing none (0%), few (<20%), many (20-50%), and most (>50%) as in Hancock and Pool (2017, p. 4). Results are presented in Table 3.1. Raters were debriefed about the inclusion of transgender communicators following data collection as they were blinded to this fact prior to making ratings.

Statistical Analyses

A multivariable prediction model of perceived masculinity-femininity ratings for transgender and cisgender communicators was developed using standard multiple linear regression. Acoustic and movement measures served as the predictor variables and mean masculinity-femininity ratings made in AV mode served as the criterion variable (i.e., DV). As in Hardy et al. (2018), purposeful selection was performed prior to the main analysis to select the set of predictor variables to be used in the multiple regression (Hosmer et al., 2013). Predictors were retained if they were significant predictors of the

criterion variable in a univariable analysis ($\alpha = .25$). Differences in perceived masculinity-femininity (DV) between presentation modes (IV) were explored using a one-way repeated measures ANOVA.

Results

Reliability of Acoustic Variables

Intra- and inter-measurer reliability were computed using Intraclass Correlation Coefficient (ICC) estimates and their 95% confidence intervals (CI) based on a two-way mixed effects model with absolute agreement. Calculations were conducted using SPSS statistical software version 24. The first author and two speech-language pathology graduate students completed repeat acoustic measurements for eight (i.e., 20%) randomly selected communicators. They were not blinded to the predictors, but did follow the same measurement protocol. First and second measures conducted by the first author were used to calculate intra-measurer reliability. Inter-measurer reliability was calculated using measures completed by the first author and the SLP students. Using criteria for 95% CI suggested by Koo and Li (2016), intra-measurer reliability was excellent (i.e., >.9) for nine of the acoustic variables and ranged from good to excellent (i.e., >.75) for ST range. Percent upward intonation shifts was not reliable and, consequently, was excluded from subsequent analyses. Inter-measurer reliability was excellent for all but two variables, ST range and SPL, which ranged from moderate to excellent reliability (i.e., >.5). Movement measures were assumed to be reliable after the movement measurement script was verified.

Reliability of Perceptual Ratings

Intra- and inter-rater reliability measures for masculinity-femininity ratings made in all three modes was assessed using ICC estimates and their 95% CI based on a two-way mixed effects model with absolute agreement. Ratings made in audio and AV modes had excellent reliability. Intra-rater reliability was poor (i.e., < .5) for ratings made in the visual mode despite good to excellent inter-rater reliability. Visual mode ratings were excluded from the statistical analysis.

Acoustic and Perceptual Ratings Data

Descriptive statistics for the communication variables and perceptual ratings are presented in Table 3.3. Mean f_o for cisgender participants was within the respective normative ranges (Baken & Orlikoff, 2000) and was at the upper end of the cisgender men normative range for transgender women. Mean values for transgender women fell between those of cisgender women and men for 75% of the communication variables and both perceptual ratings. Mean masculinity-femininity ratings in audio and AV modes for transgender women (audio = 19.22, AV = 18.44) and cisgender men (audio = 14.88, AV = 13.28) communicators fell below the midpoint of the equalized rating scale (i.e., 21.87) and thus in the masculine range. Mean ratings for cisgender women communicators (audio = 32.93, AV = 31.76) fell within the feminine range of the scale. Of note, masculinity-femininity ratings made in both the AV and audio only mode were not significantly correlated with rater sexual orientation (r_s = .043, p = .856) and (r_s = .135, p = .436) or ally status (r_s = .190, p = .423) and (r_s = .049, p = .837) in this study.

Table 3.3 *Descriptive Measures for Communication Variables and Perceptual Ratings for All Communicators and Gender Groups*

	All Communicators (n = 40)		Transgender Women (n = 20)		Cisgender Women (n = 10)		Cisgender Men (n = 10)	
Acoustic Measure	Mean (SD)	Min-Max	Mean (SD)	Min-Max	Mean (SD)	Min-Max	Mean (SD)	Min-Max
F_o (Hz)*	147.73 (40.04)	96.22-230.58	137.64 (31.53)	100.52-227.93	197.53 (20.97)	150.46-230.58	118.11 (21.21)	96.22-156-85
Average formant frequency (Hz)*	1952.22 (141.04)	1692.99-2344.64	1906.89 (82.44)	1773.17-2073.26	2124.32 (123.17)	1897.59-2344.64	1870.75 (108.78)	1692.99-2116.39
Shimmer (%)*	0.035 (0.017)	0.014-0.077	0.032 (0.014)	0.014-0.062	0.028 (0.121)	0.019-0.058	0.048 (0.021)	0.016-0.077
NHR *	0.017 (0.12)	0.002-0.137	0.012 (0.011)	0.002-0.045	0.013 (0.102)	0.003-0.034	0.033 (0.040)	0.010-0.137
SPI[a]	18.39 (8.84)	4.73-34.36	17.66 (8.38)	4.73-34.06	20.59 (11.41)	7.09-34.36	17.90 (7.94)	7.29-32.77
SPL (dBSPL)*	74.68 (3.50)	69.39-82.43	74.31 (2.69)	69.76-79.62	73.36 (3.82)	69.39-82.19	76.75 (4.04)	70.69-82.43
RoS (syllables/s)	3.70 (0.56)	2.36-4.81	3.64 (0.56)	2.60-4.81	3.90 (0.58)	2.81-4.53	3.61 (0.56)	2.46-4.31
ST range	9.0 (2.59)	4.58-16.63	8.83 (2.55)	5.17-16.63	8.33 (2.83)	4.58-14.25	10.00 (2.36)	6.25-12.63

MS Head (degrees)*	357.97	66.25-793.74	299.08 (156.06)	66.25-642.30	492.80 (182.94)	208.64-793.74	335.01 (133.62)	123.06-469.32
Hand Movement (mm)	6146.22	129.89-19,913.43	5123.59 (5370.18)	190.64-19,540.99	8039.87 (4549.85)	942.14-16,025.49	6297.84 (6373.41)	129.89-19,913.43
Palmup%	17.43	0.0-96.64	17.72 (26.01)	0.00-96.64	12.55 (14.07)	0.00-41.10	21.75 (29.94)	0.00-92.60
Elbow%	143.95	103.96-195.44	144.45 (26.16)	103.96-195.44	143.84 (13.40)	129.37-174.02	143.06 (18.09)	114.28-178.45
Masculinity-Femininity Rating AV Mode	20.48 (7.66)	9.25-34.48	18.44 (3.31)	12.54-27.14	31.76 (2.40)	26.73-34.48	13.28 (4.14)	9.25-21.41
Masculinity-Femininity Rating Audio Mode	21.51 (7.49)	10.72-35.71	19.12 (3.21)	13.16-27.41	32.93 (1.77)	29.25-35.71	14.88 (3.42)	10.72-22.01

Note. [a]SPI had missing values. Sample sizes for that variable were: all communicators = 35; transgender woman = 18; cisgender woman = 8; cisgender man = 9. The extreme outlier was excluded from MS Head. Resulting sample sizes were: all communicators = 39 and transgender women = 19. Communication variables retained through purposeful selection of covariates for use in the model-building multiple regression analysis are marked with *.

Communication-Based Predictors of Masculinity-Femininity Ratings

Purposeful selection identified six communication variables to serve as predictors for the multiple linear regression: f_o, average vowel formant, shimmer, NHR, SPL, and MS Head[4]. The results of the standard multiple linear regression (n = 39) revealed a significant model: $F(6,32) = 24.009$, $p < .001$ that explained 78.4% of the variance in masculinity-femininity ratings (adjusted $R^2 = .784$). Table 3.4 provides information about regression coefficients for the predictor variables entered into the model. Three variables emerged as significant predictors of masculinity-femininity ratings: f_o, average formant frequency, and SPL. The first two predictors were positively related to masculinity-femininity ratings. That is, as f_o and average formant frequency increased, raters perceived the communicator to be more feminine. Conversely, SPL was negatively related to the perceptual ratings; louder voices were associated with more masculine ratings.

[4] An extreme outlier was identified in MS Head. Viewing of the PLD revealed the participant's (Tcomm13) head movements were secondary to frequent upper body movement along the MS plane (i.e., leaning forward and back) rather than isolated head movements. These movements were uncharacteristic of the other participants; therefore, the outlier was excluded from the analysis for the MS Head variable only. Of note, the results were the same save for minor differences in adjusted R^2 and coefficients when the outlier was included.

Table 3.4 *Unstandardized and Standardized Regression Coefficients for the Predictors Entered into the Model*

Predictor Variable	***B***	**SE B**	β	***p*** (α = .05)
F_o	0.102[a]	0.021[a]	.534[a]	< .001*[a]
	0.113[b]	0.019[b]	.588[b]	< .001*[b]
Average formant frequency	0.017[a]	0.006[a]	.312[a]	.006*[a]
	0.020[b]	0.005[b]	.364[b]	.001*[b]
Shimmer	-69.163[a]	53.277[a]	-.155[a]	.204[a]
	-[b]	-[b]	-[b]	-[b]
NHR	11.757[a]	34.812[a]	.036[a]	.738[a]
	-[b]	-[b]	-[b]	-[b]
SPL	-0.667[a]	0.185[a]	-.304[a]	.001*[a]
	-0.539	0.163[b]	-.246[b]	.002*[b]
MS Head	.004	.004	.090[a]	.294
	-[b]	-[b]	-[b]	-[b]

Note. Significant predictors of masculinity-femininity ratings are marked with *. [a] Coefficients obtained from the first model including all five predictor variables. [b] Coefficients obtained from the second model including only the three predictors that were significant in the first model.

The regression was repeated using the three significant predictors to calculate new regression coefficients (see Table 3.4). Data from all communicators were included in the analysis (i.e., *n* = 40). The coefficients show that masculinity femininity ratings would be expected to increase by 4.5 points (.588 SD) on the equalized scale for every one SD increase in f_o (i.e., 40.04 Hz) and 2.8 points on the equalized scale (.364 SD) for every one SD increase in average formant frequency (i.e., 141.04 Hz). Ratings would be expected to drop by 1.9 points on the equalized scale for every 3.5 dB SPL increase in SPL.

The final predictive model developed for masculinity-femininity ratings of the transgender and cisgender communicators who participated in the present study is: $Y' = 5.52 + .113X_1 + .020X_2 - .539X_3$, where Y' represents the predicted value of the criterion variable (i.e., masculinity-femininity ratings made on the equalized scale), X_1 represents mean f_o, X_2 represents

average formant frequency, and X_3 represents SPL. A masculinity-femininity rating can be calculated for a given individual by inserting that individual's associated acoustic measurements in the place of the X values.

Differences in Perceptual Ratings by Presentation Mode

A one-way repeated measures ANOVA was used to identify differences in masculinity-femininity ratings as a function of presentation mode. Since ratings made in the visual mode were not reliable, the IV included two levels: audio mode and AV mode. The results revealed that communicators were perceived to be significantly more feminine in the audio mode than the AV mode $F(1, 39) = 27.22$, $p < .001$. Mean masculinity-femininity rating on the equalized scale in the audio mode was 21.51 (standard error = 1.18) compared to 20.48 (standard error = 1.21) in the audiovisual mode. Effect size estimates indicated that 41.1% of the variance in ratings could be accounted for by presentation mode (partial eta squared = .411).

Discussion

The aim of the present study was to identify a set of communication-based predictors of perceived masculinity-femininity and to explore the effect of presentation mode on these perceptions. Three acoustic variables emerged as significant predictors: f_o, average formant frequency, and SPL. Shimmer, NHR, and MS Head also were analyzed as predictors of these ratings but did not reach significance based on the present data. Masculinity-femininity ratings were significantly higher (i.e., more feminine) in the audio than the AV presentation mode and unreliable in the visual only mode.

The variables identified as significant predictors of masculinity-femininity in this study were the same acoustic variables as were identified as predictors of masculinity-femininity ratings based solely on the voice (i.e., audio only mode of presentation) in Hardy et al. (2018). These

findings add to the existing evidence for the contribution of pitch and vocal tract resonance as cues to gender attribution and provide needed preliminary evidence for vocal loudness. As such, the results endorse modifying these aspects of the voice in communication feminization or masculinization training. In addition, once validated, the predictive model may be used in the clinical context to predict how listeners will perceive a client's voice. These predictions could be used as baseline measures and to track progress in training.

Speaking frequency and vocal tract resonance modification already are common goals for feminization training (Davies et al., 2015; Hancock & Garabedian, 2013) and this practice should continue. Perhaps, however, more emphasis could be placed on SPL modification. This type of adjustment may be particularly helpful for transmasculine individuals who have not achieved desired results with or have opted not to pursue hormone replacement therapy. Speaking with increased SPL may help to further masculinize the voice. Although voice and communication modification training historically has focused on feminization with transgender women, many transmasculine individuals are seeking SLP assistance to masculinize their communication (Davies et al., 2015). Future research should explore the effect of modifying SPL on changing perceptual ratings related to gender attribution for transgender and GD individuals. Research also is needed to determine whether the predictors of these perceptual ratings are the same for transmasculine communicators as there is a paucity of studies including these participants and training should be informed by relevant research.

The results of the purposeful selection indicated head movements were somewhat predictive of perceptual ratings on their own, but not strong enough cues of masculinity-femininity when considered with the voice. The relative salience of head movements and the positive relationship

with perceptual ratings is in keeping with the results of Hill and Johnston (2001) who found that rigid head movements provided cues to identify a speaker's sex. In that study, greater use of head movements was associated with more feminine ratings.

It was unexpected that no motion variables emerged as significant predictors in the overall model in the present study, given the findings of Hill and Johnston (2001) and other evidence in the cisgender research literature for gender differences in amount of gesture use (Hall, 1984; Hostetter & Hopkins, 2002; Nicoladis et al., 2007). Four factors may help to explain this exclusion. First, it is possible that the simple nonverbal behaviors measured in the present study do not convey meaningful information about gender when isolated from appearance. This explanation seems unlikely given the beliefs about nonverbal communication behaviors that are held in Western culture. In general, women are believed to use more expressive nonverbal behaviors than men and, specifically, are believed to use more hand gestures (Briton & Hall, 1995). Moreover, it has been reported that transgender women believed they naturally had feminine body language (Pasricha et al., 2008), which suggests that certain behaviors are associated with masculine or feminine presentations and attributions. Comments made by several raters further endorse these relationships. Those raters reported that frequent use of hand gestures made the communicator seem more feminine, although this effect seemed to vary with presence or absence of the voice in the recording.

A second explanation is that some communicators may have experienced a Hawthorne effect during the cartoon description task recording resulting in more constrained use of gestures than would be typical for their communication style. For example, several participants were observed to gesture frequently with their hands during casual conversation, but refrained from doing so during the recording task. Participants were asked to retell the story of the cartoon as if the first

author had not seen it before or as if telling it to a friend. They were not provided with explicit instructions about their gestures in an effort to elicit behaviors that were as natural as possible. Although necessary, this lack of instruction did not redirect participants who believed they were not supposed to move because they were being recorded. Movements also may have been influenced by the positioning of the first author relative to the communicators (i.e., lower and slightly to the side) and possible perseveration on the calibration task (i.e., looking at the dot on the wall and sitting up very straight rather than looking at and engaging with the first author).

The third and fourth explanations are factors related to the raters. It is possible they were paying attention to movements or body positions not measured in this study and/or were engaging in "bias checking" when making their ratings. For example, even though no markers were placed on the lower body of the communicators, cues about the position of the legs may have been inferred or assumed based on the position of the hands if they were resting on the lap. One rater indeed reported noticing "man spread" based on hand position. It should be noted that the same or similar differences in distance between the hands could be achieved with the legs together if the hands were resting on the sides versus the tops of the thighs. Another rater stated that he found the hands clasped in the lap to be feminine and leaning and pointing gestures (pantomimed to be imposing in nature) to be masculine. These comments suggest that the role of gestures and body position in conveying masculinity-femininity requires further investigation.

Finally, it appeared that raters were more willing to use gender normative assumptions to make ratings of the voice than of gestures and body position. Several raters (all men) described initially using gestural information to make their ratings (i.e., assigning more feminine ratings to communicators who used hand gestures), but then realized they too used similar gestures and thus became aware of their inherent stereotypes and attempted to refrain from using this

information as a cue. An extreme example was R14 who stated he knew he couldn't determine someone's gender based on movements and therefore gave every communicator the same masculinity-femininity rating. This self-monitoring may have masked the true predictive value of gestures on perceived masculinity-femininity in every day interactions.

"Bias checking" also may have contributed to the lack of reliability of ratings in the visual presentation mode that resulted in their exclusion from this study. More explicit instructions about basing ratings on first instincts and affirming the acceptability of basing ratings on previously held beliefs about communication may have improved reliability, revealed a greater contribution of nonverbal communication to perceptual ratings, and better represented the automatic attributions often made during social interactions.

Despite excluding visual mode ratings, results of this study showed a significant effect for mode of presentation (i.e., ratings were significantly higher, or more feminine in the audio only mode versus AV mode). Visual information moderated ratings toward masculine. These results are inconsistent with previous research, which found ratings of femininity or femaleness to be highest for transgender women speakers and avatars in the visual mode, followed by audiovisual mode, and finally, audio mode (Boker et al., 2011; Van Borsel et al., 2001). In these studies, stimuli that included visual information increased perceived femininity. It should be noted that these stimuli also showed the communicator's appearance (real or several avatars rendered to look like the communicator and peers of the same and opposite sex), including facial structure; however, Boker and colleagues (2011) were able to isolate the effect of head movements from appearance. They found that femininity ratings were based on a combination of voice and motion information whereas masculinity ratings predominantly were based on the voice. The mean masculinity-femininity rating for transgender women and cisgender men communicators in the

present study fell within the masculine range of the equalized scale (i.e., below the midpoint of 21.87); therefore, raters may have been working with a masculinity scale for many of the communicators and thus, basing ratings primarily on the voice even when gestures were present.

Other factors such as bias checking and inferring information from unmeasured movements and positions discussed in the previous section also may help to explain the masculinizing effect of visual information in this study. It also is possible that gender-based differences in gesture use existed. Although transgender women have reported feeling that feminine nonverbal behaviors come naturally to them (Pasricha et al., 2008), this has not yet been measured systematically. Given that many of these individuals were socialized in masculine gender roles, it is possible that they, generally speaking, use masculine body language. The more masculine ratings in AV mode as compared to audio only mode would then make sense, as the majority of communicators in this study would be expected to have masculine nonverbal behaviors. The majority of transgender participants in this study (60%) had not yet received voice and communication feminization training and masculine behaviors may have persisted beyond their gender role transition.

A final explanation for the relatively masculine ratings in AV mode is the effect of stereotype-based expectancies on social perception. Individuals tend to rely more heavily on stereotypic expectancies when perceptual information provides strong cues to membership in a social group (Hamilton et al., 1990). The acoustic information in the voice signal served as a strong cue to masculinity femininity and gender attribution [as reported in Hardy et al., 2018)] for the communicator group. Given that the majority of the communicators were attributed as men (Hardy, et al., 2018), raters may have then relied on masculine stereotypic expectancies when considering the additional information (i.e., gestures) provided in AV mode and consequently

assigned more masculine ratings. Indeed, several raters who were presented with the visual only stimuli in the first rating block commented that they wished they had heard the voice in the first block so they could "figure out the gender" of the communicator before rating gestures.

Taken together, the results of this study suggest the voice provides stronger cues to masculinity-femininity than gestures. Gestures do, however, appear to play a role in conveying masculinity-femininity to some extent, possibly as part of stereotypic beliefs as described above. More research is needed to illuminate this role and should endeavor to include communicators across the continuum of gender positioning, expression, and attribution as well as investigate differences in gesture use between gender groups and attributed gender groups. Future researchers also may wish to use animations or avatars (created from motion data) as perceptual stimuli rather than PLDs in order to explore more nuanced aspects of gesture. Based on the present findings, clinicians should be aware of the potential contribution of body language to perceived masculinity-femininity and target these behaviors with the understanding that their contribution to overall gender attribution likely is minimal.

Study Limitations

In addition to the limitations discussed in the previous section with respect to data collection procedures, variable selection, and rater bias, the present study also was limited by a relatively small sample size and lack of variability in participant characteristics. Although the sample size was large when compared to other communication studies including transgender participants, the case-to-predictor ratio was small. It is possible that different predictors may have emerged as significant (e.g., gestural variables) had there been a larger sample size and, thus, greater power. Given the relatively small population of transgender people, multi-site collaborations may be needed to recruit an adequate number of participants.

Similarly, there was an underrepresentation of feminine communication characteristics in the present study as evidenced by the masculinity-femininity ratings obtained by the gender groups. Transgender women made up half of the total sample and generally were perceived to communicate in a masculine fashion. The mean masculinity-femininity rating for the transgender women was between those of the cisgender groups, but markedly closer to ratings assigned to the cisgender men. It would be valuable to repeat this study with more even participant representation across the masculinity-femininity spectrum to ensure the developed model adequately predicts the lived experience of transgender women. Moreover, the present study did not include transgender men. As more of these individuals are seeking voice and communication modification services, it is becoming increasingly important that the evidence base informing care is expanded accordingly. Future studies should explore relationships between communication characteristics and perceptual ratings amongst transgender men, specifically, as well as communicators representing the spectrum of gender diversity to obtain a more complete understanding of the ways in which communication contributes to gender attribution for transgender people.

Finally, the communicators in this study were English speakers predominantly of European ancestry. Consequently, generalization of results is limited to individuals with similar personal characteristics. Future research should consider intersectionalities between gender, language, and culture to further elucidate the role of communication in gender attribution.

Conclusions

Both aims of this study were met: three acoustic variables were identified as predictors of masculinity-femininity ratings and mode of presentation was found to have a significant effect on these ratings. Results confirmed the salience of voice in conveying aspects of one's gender to

others and underscore the vital role of communication modification training for those for whom correct gender attribution is important. Clinicians should continue to focus on helping clients modify speaking f_o and vocal tract resonance and consider placing greater emphasis on vocal loudness, especially for those clients who wish to sound more masculine. This was the first study to investigate nonverbal communication amongst transgender communicators and results provided preliminary evidence that nonverbal behaviors make a small contribution to perceptual measures. This line of inquiry warrants further investigation.

CHAPTER 4: ASSOCIATIONS BETWEEN VOICE AND GESTURAL CHARACTERISTICS OF TRANSGENDER WOMEN AND SELF-RATED FEMININITY, SATISFACTION, AND QUALITY OF LIFE

Abstract

Purpose: Client-based subjective ratings of treatment and outcomes are becoming increasingly important as speech-language pathologists embrace client-centered care practices. Of particular interest, is the value in understanding how these ratings are related to aspects of voice and communication modification training programs for transgender and gender diverse individuals. The purpose of the present observational study was to explore relationships between acoustic and gestural communication variables and communicator-rated subjective measures of femininity, communication satisfaction, and quality of life (QoL).

Method: Twelve acoustic and gestural variables were measured from high-fidelity audio and motion capture recordings of transgender women (n=20) retelling the story of a short cartoon. The participants also completed a set of subjective ratings using a series of likert-type rating scales, a generic QoL questionnaire, and a population-specific voice-related QoL questionnaire. Correlational analyses were used to identify relationships between the communication measures and subjective ratings.

Results: Significant negative relationships were identified between use of palm-up hand gestures and ratings of gestural femininity and satisfaction with overall communication. The acoustic variable of average semitone range was positively correlated with overall QoL. No acoustic measures were significantly correlated with voice-related QoL and unlike previous studies; speaking f_o was not associated with any of the subjective ratings.

Conclusions: The results from the present study suggest that voice characteristics may have limited association with communicator-rated subjective measures of communication satisfaction or QoL for this population. Results also provide preliminary evidence for the importance of nonverbal communication targets in voice and communication modification training programs.

Introduction

Voice and communication feminization training, as provided by a skilled and culturally sensitive speech-language pathologist (SLP), represents an essential aspect of gender affirming care for many transgender and gender diverse (GD) people (Adler, Hirsch, & Mordaunt, 2012; Bodoin, Byrd, & Adler, 2014; Davies, Papp, & Antoni, 2015). The goal of such training is to modify communication behaviors (i.e., aspects of voice, speech, language, and nonverbal communication) to better align with gender positioning (i.e., one's subjective sense of their gender) and, ultimately, improve quality of life (QoL) (ASHA scope of practice for SLP, 2016). More specifically, communication behaviors that are incongruent with gender positioning and socially-constructed expectations for gendered behavior frequently result in distressing experiences such as misattribution by others (Dacakis et al., 2013), social participation restrictions (Byrne, 2007; Hancock, 2017; Hardy et al., 2013), and increased risk for discrimination and violence (Bauer & Scheim, 2015; James et al., 2016a). It follows that communication behaviors eliciting correct gender attribution can potentially positively impact QoL. Moreover, transgender women have reported greater communicative satisfaction with interactions when their gender is accurately attributed (Pasricha et al., 2008).

In line with emerging client-centered care practices, SLPs increasingly are realizing the importance of the client's own perception of and satisfaction with their communication in driving the training process. Accordingly, client-perceived changes in voice and QoL were

identified as important factors in determining when to discharge from traditional voice therapy (Gillespie & Gartner-schmidt, 2018). These same client-generated elements could potentially be used for guiding voice and communication feminization training (Pasricha et al., 2008). For example, self-ratings of communication made by transgender women have been more strongly associated with QoL than ratings made by others (Hancock et al., 2011; E. J. M. McNeill et al., 2008), further supporting their value in the context of goal-setting and discharge planning. For this reason, it is important to understand the relationships between these client-rated subjective measures of communication and QoL and aspects of communication such as speaking fundamental frequency f_o that might serve as relevant training targets and produce meaningful outcomes. Advancing our understanding these relationships is the focus of the present study.

Relationships Between Communication and Client-Rated Subjective Measures of Communication and Quality of Life

A few studies have begun to explore relationships between communication characteristics and subjective ratings of communication and QoL made by transgender women; however, this area remains underexplored. Communication factors of interest in existing studies have been limited to f_o in most cases and otherwise have only included a small set of other acoustic variables (e.g., jitter). Associations with many aspects of communication have yet to be fully investigated including those commonly targeted during training (e.g., measures of vocal tract resonance), known and/or suspected predictors of gender attribution (e.g., sound pressure level [SPL]), or other non-acoustic aspects of communication (e.g., nonverbal communication).

F_o has been significantly positively correlated with self-rated masculinity-femininity of the voice (E. J. M. McNeill et al., 2008; Owen & Hancock, 2010) as well as satisfaction with voice (Holmberg et al., 2010) and overall communication (Byrne et al., 2003). Self-rated masculinity-

femininity also was negatively correlated with semitone (ST) range in one study (Owen & Hancock, 2010). Another study demonstrated improved self-rated vocal femininity and satisfaction with voice following oral resonance training that resulted in raised vowel formant frequencies; however, it was not determined whether or not these variables were significantly related (Carew et al., 2007). In each case, ratings were made using a visual analog scale (VAS) and acoustic measures were obtained from connected speech recordings. Despite these promising findings, the research has not been conclusive with respect to a relationship between f_o and voice satisfaction (E. J. M. McNeill et al., 2008). Moreover, f_o accounted for only 18% of the variability in communication satisfaction for the transgender women (n = 21) who participated in Byrne, Dacakis, and Douglas's study (2003). These results suggest that variables other than f_o contribute to communicative success and client well-being.

Pasricha, Dacakis, and Oates (2008) identified five categories that were important to transgender women with respect to their communication satisfaction. Among these was "Features of Communication", which included voice, vocabulary, body language, speech, and intonation. In their study, participants indicated they were least satisfied with their voice and felt it was the aspect of their presentation that was most likely to lead to misattribution. They also reported their nonverbal communication was naturally feminine. These findings suggest links between both acoustic and nonverbal communication measures with perceptual outcomes. For example, a low f_o or vowel formant frequencies (VFF) may be related to negative outcomes whereas using more hand gestures may be related to positive ones. Based on the Pasricha et al., study, multiple aspects of communication should be considered and further support the need to advance our understanding about communication factors other than f_o, and their relationships to communication satisfaction.

Taken together, the available research shows promising preliminary evidence for associations between communication measures and self-rated femininity and communication satisfaction. These relationships also may extend to QoL as several studies have shown significant correlations between self-rated femininity and/or voice satisfaction and voice-related QoL (VRQoL) (Dacakis et al., 2017a; Hancock et al., 2011; E. J. M. McNeill et al., 2008): higher self-ratings were associated with higher VRQoL.

Only one study has explored relationships between communication variables and measures of QoL (Dacakis et al., 2017a). Those authors found no significant correlations between f_o, jitter, shimmer, and noise to harmonic ratio (NHR) and scores on the Transsexual Voice Questionnaire (TVQ^{MtF}) (Dacakis & Davies, 2012), which is a population-specific questionnaire that measures the impact of the voice on daily life. The acoustic variables included in their study do not fully represent the features of communication identified as important to communication satisfaction in Pasricha et al. (2008) or those that are thought to cue gender attribution, and as such, further investigation is warranted (Dacakis et al., 2017a; Hancock et al., 2011).

Measures of VRQoL offer a limited view of QoL as they consider only communication aspects and not other components important to QoL. QoL may comprise elements such as personal relationships, stress levels, and physical health, among others (Ferrans & Powers, 1985). As stated previously, communication, life participation, and personal safety are interrelated for many transgender and gender diverse people and all, therefore, may influence QoL. For this reason, a measure of overall QoL offers a more complete picture of a person's well-being than VRQoL. Indeed, Cruice, Worrall, and Hickson (2000) recommended using both

generic and condition-specific QoL tools as outcome measures in SLP practice. To date, no studies have explored relationships between communication variables and overall QoL amongst transgender women.

Study Aims

The present study aims to identify relationships between an array of communication variables and communicator-rated subjective measures related to perceived femininity, communication satisfaction, and QoL. Specific questions related to these objectives include: 1) How are acoustic measures related to communicator-rated a) voice femininity?, b) voice satisfaction?, c) VRQoL?; 2) How are gestural measures related to communicator-rated gestural femininity?; 3) How are acoustic and gestural communication measures related to communicator-rated a) communication femininity?, b) communication satisfaction; and c) overall QoL?

Methods

The present observational study was approved by the Human Research Ethics Board (HREB) at the University of Alberta (Pro00060133) prior to participant recruitment. It also received operational approval from Alberta Health Services and Covenant Health. STROBE reporting guidelines (von Elm et al., 2007) were adopted for the present manuscript.

Participants

A group of 20 transgender women participated in the present study. This sample size was similar to or larger than other communication studies including transgender participants and represented the number of participants that could reasonably be expected to be recruited from the Edmonton area. All participants had been living in the feminine gender role the majority of the time (i.e., $\geq$ 80%) for at least six months. These criteria were included to ensure participants had adequate experience in their affirmed gender role to respond to the items on the instruments used

to measure QoL. Participants were recruited from gender clinics, a voice and communication feminization training program, LGBTQ2S support centres, and the Edmonton region using convenience sampling, recruitment letters, and a media release. The mean age of the participants was 41.2 (SD = 14.38) years. They had been living fulltime in the feminine gender role on average 7.6 years (SD = 11.72), but this value ranged from nine months to 52 years. Nearly all participants (i.e., *n* = 18) were receiving hormone replacement therapy, but only 40% (*n* = 8) had accessed voice and communication feminization training or completed gender affirmation surgery (GAS). Several other participants were awaiting access to these services.

Procedures

Data collection took place in the Syncrude Centre for Motion and Balance at the Glenrose Rehabilitation Hospital in Edmonton between June 2016 and May 2017. After consenting to participate in the study, participants provided demographic information and then proceeded with completing the subjective ratings of femininity, communication satisfaction, and QoL. Acoustic and motion capture recordings followed.

Communicator-rated femininity and communication satisfaction. Participants rated the femininity of their own communication (i.e., voice, gestures, and overall communication) as well as their voice and communication satisfaction by answering a series of five questions using a seven-point rating scale. Each scale was bounded with "Not at all feminine/satisfied" and "Very feminine/satisfied" and marked with "Somewhat feminine/satisfied" at the midpoint.

Quality of life measures. The TVQ^{MtF} (Dacakis & Davies, 2012) was selected to measure VRQoL in the present study. This tool was designed specifically for transgender women, has been translated into more than 10 languages, and has a growing body of evidence demonstrating its reliability and validity (Dacakis et al., 2016, 2013, 2017a; Dacakis, Oates, & Douglas, 2017b;

Dacakis, Pasricha, & Oates, 2005; Davies & Johnston, 2015). The questionnaire consists of 30 items related to vocal functioning and social participation. Individuals rate their perceived experience with each item on a four-point scale, where lower numbers represent lower frequency of occurrence. The lowest possible total score is 30 and the highest is 120. Minimum and maximum scores for the vocal functioning and social participation components are 14-56 and 12-48, respectively (Dacakis et al., 2017c). Higher scores represent greater negative impact of the voice on daily life.

The Quality of Life Index (Generic Version-III) (Ferrans & Powers, 1998) (QLI) was used to measure overall QoL in the present study. The instrument consists of two sections: satisfaction and importance. In the first section, individuals rate their satisfaction with 33 aspects of their life (e.g., education, relationships, amount of worry, health) on a six-point scale where one represents "Very dissatisfied" and six represents "Very satisfied". In the next section, individuals rate the relative importance of each of these areas on a similar six-point scale anchored with "Very unimportant" and "Very important". Overall QoL scores as well as scores for four subscales (i.e., Health and Functioning, Social and Economic, Psychological/Spiritual, and Family) are available; however, only the overall score was used for the present study. All scores are normalized and weighted according to rated importance during the scoring process. The QLI has been translated into multiple languages and multiple condition-specific versions have been developed. It is reliable, valid, and sensitive to change as evidenced by more than 48 studies listed on the QLI website (https://qli.org.uic.edu/reliability/reliabilityhome.htm).

Communication Measures. Simultaneous acoustic and motion capture recordings of participants' voices and gestures were obtained during a cartoon retell task of the Pink Panther cartoon "In the Pink of the Night" (Davis, 1969). Additional acoustic recordings captured the

participants performing sustained vowel (i.e., /a/ for three to five seconds, five repetitions) and carrier phrase reading (i.e., "say hVd again" using the vowels /i, a, u/, five repetitions) tasks. Full descriptions of acoustic and gestural measurement procedures can be found in Hardy et al. (2018) and Hardy et al. (under review, *Journal of Speech, Language, and Hearing Research*), respectively.

A total of 12 communication variables were included in the present study, eight acoustic and four gestural. These variables were selected based on a review of the current voice and communication modification literature as well as research investigating gender differences in communication behaviours or relationships between communication behaviours and ratings related to gender attribution. The reader is directed to Davies et al. (2015), Leung, Oates, & Pang Chan (2018), and Hardy et al. (under review, *Journal of Speech, Language, and Hearing Research*) for relevant reviews.

Acoustic variables included: mean f_o (Hz); average formant frequency for the vowel /i/ (i.e., an average of the first three formants) (Hz); mean rate of speech (RoS) (syllables/s); mean SPL (dBSPL); mean semitone (ST) range; mean NHR; mean shimmer (%); and mean soft phonation index (SPI). Vowel formant frequencies were measured at the midpoint of the vowel in the carrier phrase "say /hid/ again" and averaged across the five productions. Voice perturbation and periodicity variables were measured during the middle one second of the sustained vowel and averaged across the five productions. All other acoustic variables were measured within a 30-45 second sample of the cartoon retell acoustic recording that was balanced for linguistic complexity across participants. Each variable first was measured within the individual t-units

that comprised the sample and then averaged across t-units. A t-unit is "one main clause plus whatever subordinate clauses happen to be attached to or embedded within it" (Hunt, 1965, p. 305).

Gestural variables included: total head movement in degrees along the mid-sagittal plane (MS Head); total movement in millimeters (mm) of the lead hand (i.e., the hand that moved most) irrespective of direction (Hand Mvt); percent of time the lead hand was in the palm-up versus palm-down position (Palmup%); mean percent elbow: shoulder width (Elbow%). Head and hand movement variables represented nodding and amount of hand gesturing, respectively. All gestural variables were measured within the motion capture recordings of the cartoon retell during the same 30-45 second sample as the acoustic variables.

Statistical analyses. Correlational analyses were used to assess relationships between the communication and communicator-rated variables. Pearson correlation was used for the two analyses involving QoL scores (i.e., the total scores from the TVQ^{MtF} and the QLI), whereas Spearman correlation was used for the remaining analyses involving femininity and communication satisfaction ratings.

Results

Relationships with Communicator-Rated Femininity

Results from a series of Spearman correlational analyses revealed one significant relationship ($\alpha = .05$) between communication variables and communicator-rated femininity: palmup% was negatively correlated with communicator-rated gestural femininity ($r_s = -.462$, $p = .040$). The more time participants spent with their lead hand in the palm-up position, the more masculine they rated their gestures and the more time spent in the palm-down position, the more feminine they rated their gestures (see Figure 4.1). The relationship with hand position accounted for

25.1% of the variability in gestural femininity ratings (R^2 = .251). No significant relationships existed between any communication variables and communicator-rated communication femininity or between any acoustic measures and communicator-rated voice femininity. A complete listing of the correlation results is presented in Table 4.1.

Figure 4.1. Relationship Between Hand Position and Gestural Femininity Ratings

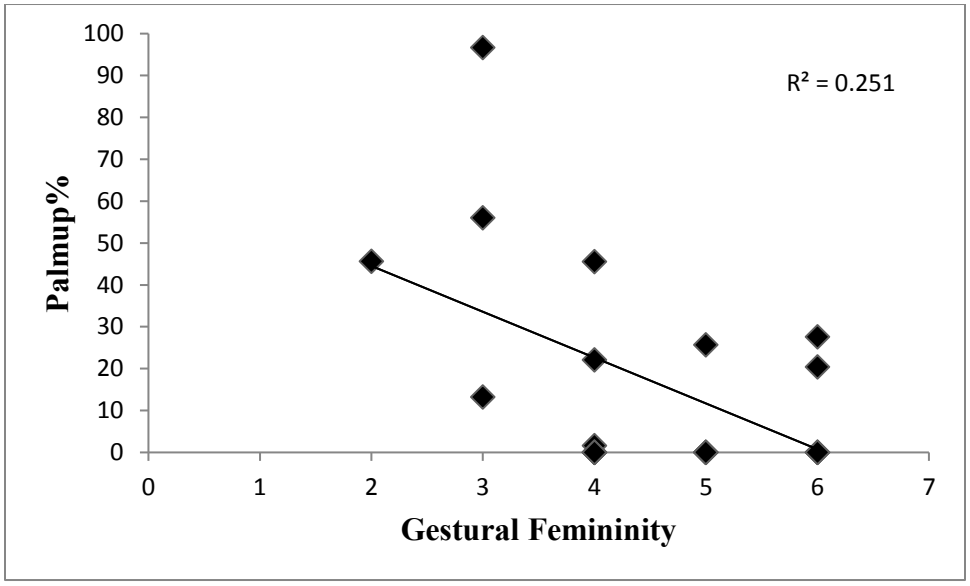


Figure 4.1. The significant relationship between hand position during a cartoon retell task and communicator-rated gestural femininity.

Table 4.1 *Correlations Between Communication Variables and Communicator-Rated Femininity, Satisfaction, and QoL*

	Communicator-Reported Ratings										QoL Measures			
Communication Variable	Comm. Femininity		Voice Femininity		Gesture Femininity		Comm. Satisfaction		Voice Satisfaction		QLI Total Score		TVQ^{MtF} Total Score	
	r_s	*p*	r_s	*p*	r_s	*p*	r_s	*p*	r_s	*p*	*r*	*p*	*r*	*p*
F_o	.089	.708	-.022	.926	-	-	.019	.936	.059	.803	-.156	.512	.021	.930
Avg Formant Freq	-.021	.931	-.119	.617	-	-	-.037	.876	-.071	.766	.041	.865	-.026	.912
RoS	-.100	.675	-.275	.241	-	-	-.082	.730	-.219	.354	-.210	.375	.122	.610
SPL	.026	.913	-.024	.920	-	-	-.063	.793	-.059	.803	-.053	.824	-.026	.913
ST range	.103	.666	.149	.532	-	-	.013	.957	.102	.668	.463	**.040***	-.199	.401
NHR	-.151	.526	-.045	.852	-	-	.041	.863	.024	.919	-.018	.939	.189	.425
Shimmer	-.129	.587	-.067	.779	-	-	.032	.893	-.010	.967	.044	.854	.164	.489
SPI (*n* = 18)	-.093	.713	.073	.773	-	-	.147	.561	.284	.254	.362	.140	-.184	.465
MS Head (*n* = 19)	.010	.968	-	-	.122	.608	-.083	.736	-	-	.186	.445	-	-
Hand Mvt	-.096	.687	-	-	-.006	.979	-.278	.235	-	-	-.014	.953	-	-
Palmup%	-.301	.197	-	-	-.462	**.040***	-.572	**.008***	-	-	.224	.342	-	-
Elbow%	.246	.296	-	-	.002	.995	.057	.811	-	-	-.330	.155	-	-

Note. Comm. = Communication; Avg Formant Freq = Average Formant Frequency; an extreme outlier (tcomm13) was excluded from MS Head resulting in *n* = 19; there were two missing values for SPI due to error results in the acoustic analyses resulting in *n* = 18; statistically significant correlations are indicated with *****.

Relationships with Communicator-Perceived Satisfaction

Palmup% also was significantly negatively correlated with communicator-rated satisfaction with overall communication (r_s = -.572, p = .008) (see Figure 4.2). Positioning the hand with the palm up was associated with lower communication satisfaction ratings for this set of participants and 29.99% of the variability in satisfaction ratings could be accounted for by this relationship. No significant relationships were revealed between any acoustic variables and communicator-perceived satisfaction with voice (see Table 4.1).

Figure 4.2. Relationship Between Hand Position and Communication Satisfaction

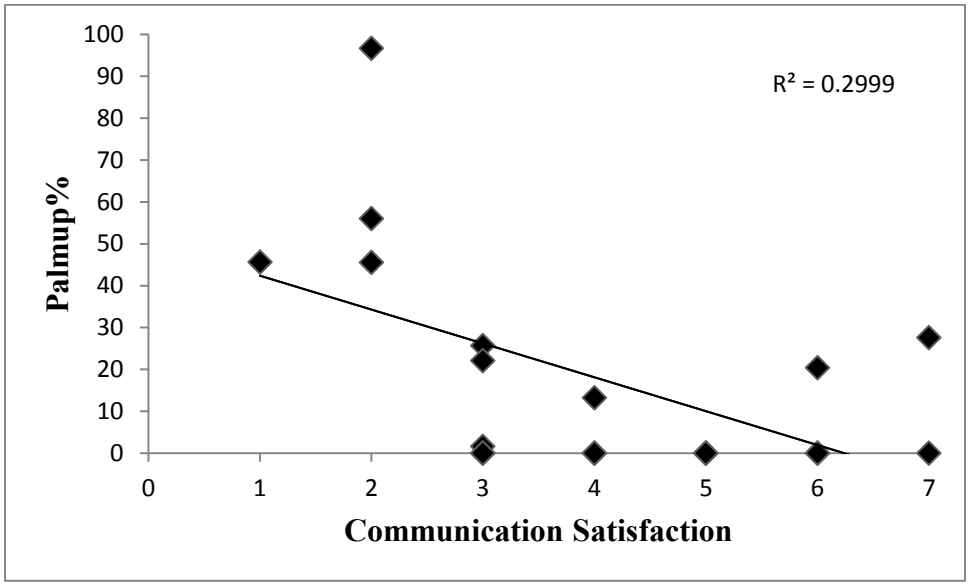


Figure 4.2. Relationship between hand position during a cartoon retell task and communicator-rated satisfaction with overall communication.

Relationships with Quality of Life

A series of bivariate Pearson correlations revealed a significant positive correlation between the overall QoL score from the QLI and mean ST range (r = .463, p = .040). That is, the greater the extent of intonation contours a participant used, the higher their overall QoL rating (see Figure 4.3). The relationship with mean ST range accounted for 21% of the variability in overall

QoL scores (R^2 = .21). Initially, a significant positive correlation also was identified between the QoL rating and MS Head (r = .450, p = .046); however, that was only the case when an extreme outlier (tcomm13) was included in the analysis. Visual inspection of the motion recording revealed that head movements for this participant were due to frequent forward and backward leaning of the upper body rather than isolated vertical head movements. This participant's MS Head data subsequently were excluded since they did not represent the variable of interest or the movements of the other participants. None of the acoustic variables was significantly correlated with VRQoL as measured by the total score on the TVQ^{MtF} or the sub-factors of vocal functioning and social participation.

Figure 4.3. Relationship Between ST Range and Quality of Life

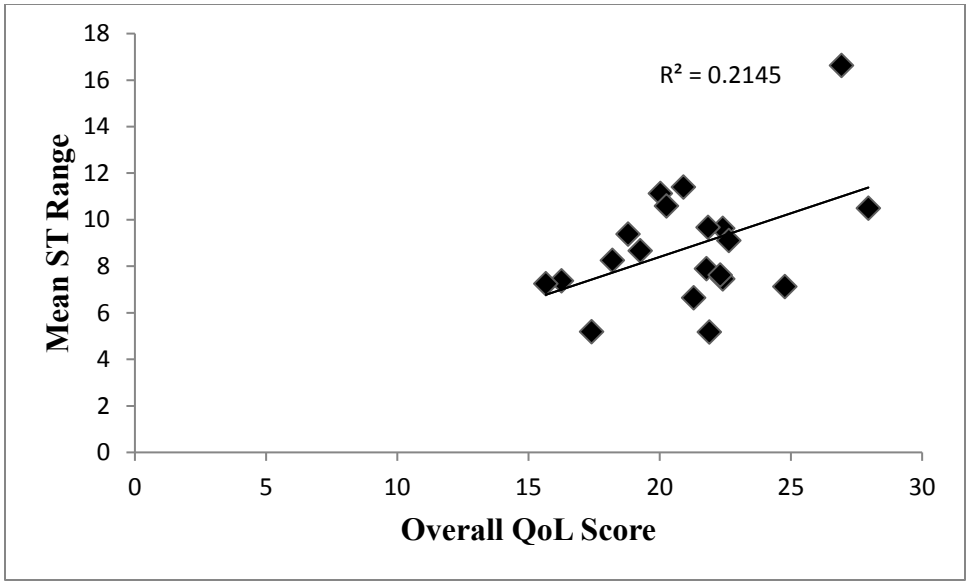


Figure 4.3. Relationship between the mean ST range used across utterances in a cartoon retell task and overall QoL as measured by the QLI (Generic Version – III) (Ferrans & Powers, 1998).

Discussion

This study aimed to explore relationships between a collection of voice, speech, and gestural communication variables and measures of communicator-rated femininity, satisfaction, and QoL amongst a group of transgender women. Correlational analyses revealed three significant relationships between communication variables and QoL measures. First, higher use of palm-up hand gestures was related to lower ratings of gestural femininity. Second, higher use of palm-up hand gestures was related to lower ratings of satisfaction with overall communication. Third, greater mean ST range was related to higher ratings of overall QoL.

To the best of our knowledge, gestural communication has been systematically measured with transgender women in only one previous study (Hardy et al., under review, *Journal of Speech, Language, and Hearing Research*). The novel findings in the present study of significant relationships with femininity and satisfaction ratings, however, speak to the value placed on nonverbal communication by these communicators. Moreover, these findings are consistent with existing qualitative research in which a group of transgender women (n = 12) reported being satisfied with their body language because they felt it was naturally feminine (Pasricha et al., 2008). Participants in the present study who tended to position their hands with the palms oriented downward also were more satisfied with their communication and perceived their gestures to be more feminine. Such a position can be achieved when the hands are resting in the lap and it is possible that the present participants found this to be a feminine posture. Indeed, in one study (Peterson, 1976), this behavior was used by cisgender women participants only. This explanation is plausible for some participants in the present study who moved their hands very little during the retell task. Of the 10 participants (50%) who always had their lead hand in the palm-down position, seven moved their hand less than 100 cm over the duration of the 30-45

second description, suggesting that it often was stationary. For those participants, it may have been the specific hand posturing and/or associated lack of movement that was related to the subjective ratings rather than a general palm-down hand orientation. Participants were not asked to explain their ratings; therefore, it is not possible to confirm whether or not this was the case. Moreover, the explanation does not account for the behaviors of all participants as some moved their lead hand a good deal (i.e., up to 740 cm) with the hand palm-down for the entirety of the sample, suggesting gestures also were made using this hand orientation. These findings are contradictory to a previous study that found cisgender women use more palm-up gestures than men (Peterson, 1976), if one is to assume that such gestures also are believed or perceived to be feminine. Whereas the present authors are not aware of any research to support such claims, there is evidence to suggest palm-down hand positioning (e.g., "limp wrist") is believed to be a feminine or effeminate characteristic, at least for homosexual men (Sánchez, Greenberg, Liu, & Vilain, 2009) and may reflect a more pervasive North American stereotypical belief.

The positive relationship between mean ST range and overall QoL ratings is, in some ways, more straightforward to understand. In the present study, individuals who had a better QoL used a wider pitch range when describing the cartoon. That is, they used greater variability in their intonation. When a person is happy, their communication may be more animated whereas when they are sad, dissatisfied, or depressed their mannerisms may be more subdued. In fact, the intonation patterns of depressed individuals have been described as monotonous and pitch variability and intonation have been negatively related to severity of depression in both children and adults (Kazdin, Sherick, Esveldt-Dawson, & Rancurello, 1985; Mundt, Snyder, Cannizzaro, Chappie, & Geralts, 2007; Nilsonne, Sundberg, Ternström, & Askenfelt, 1988). Participants in the present study, who obtained low scores on the QLI, would have rated themselves as

dissatisfied with aspects of their lives they considered important (e.g., relationships, job, amount of worry, health). When one is dissatisfied with multiple important areas of their life, they may feel a general sense of unhappiness[5]. Whereas dissatisfaction or unhappiness (i.e., low QoL) are not necessarily indicative of the presence of depression, they may manifest in similar communication behaviors. Reported negative correlations between QoL and depression across a variety of populations support the plausibility of such similarities. The fact that ST range was not also related to femininity or satisfaction ratings further supports this explanation. Information about current or previous mental health was not obtained from the study participants so it was not possible to elucidate the role of potential comorbid depression. Given that the prevalence of depression and/or psychological distress is high amongst transgender and GD people (Bauer & Scheim, 2015; James et al., 2016a) it would be valuable to explore this area in the future.

The lack of relationships between various communication variables and the subjective ratings in the present study was unexpected. For example, frequent use of hand gestures is widely believed to be a feminine nonverbal communication characteristic and has been identified by transgender women as something they do naturally (Pasricha et al., 2008), yet it was not significantly correlated with any rating of femininity, satisfaction, or QoL in this study. Similarly, previously-identified relationships between f_o and self-rated femininity of the voice as well as voice and communication satisfaction (Byrne et al., 2003; Holmberg et al., 2010; E. J. M. McNeill et al., 2008; Owen & Hancock, 2010) were not replicated in this study. These results were particularly surprising given the strength of f_o as a cue to gender attribution (Gelfer & Mikos, 2005; Gelfer & Schofield, 2000; Hardy et al., 2018; Hardy et al., under review, *Journal*

[5] The QLI included items in which participants rated their satisfaction with happiness in general and life in general. Responses to individual items were not explored in the present study; however, these ratings were represented in the total score.

of Speech, Language, and Hearing Research; Holmberg et al., 2010; King, Brown, & McCrea, 2012; Leung et al., 2018) and rater-perceived femininity (Hardy et al., 2016; Hardy et al., 2018; Hardy et al., under review; E. J. M. McNeill et al., 2008; Owen & Hancock, 2010). Since correct attribution has been related to high levels communication satisfaction for transgender women (Pasricha et al., 2008), one would expect a similar relationship to exist between communication satisfaction and f_o. The lack of relationship between f_o (or any other acoustic variable) and VRQoL, however, confirms the findings of Dacakis and colleagues (2017a) and suggests that factors other than acoustic characteristics are at play.

Perhaps these findings reflect a general self-acceptance, confidence, and/or sense of being feminine that override acoustic characteristics for some participants, with positive or negative consequence. One participant (tcomm1), for example, rated her voice as very feminine (i.e., 6 out of 7), yet her f_o was well within the normative range for cisgender men (i.e., 100 Hz) (Baken & Orlikoff, 2000). She was eight years post-transition, six years post-voice and communication feminization training, and had completed GAS. Conversely, two other participants whose f_o was within the normative range for cisgender women (i.e., 228 Hz and 196 Hz) (Baken & Orlikoff, 2000) rated their voices as very masculine (i.e., 2 out of 7). One of these participants (tcomm3) was more than two decades post-transition, had completed GAS, and was more than 10 years post-voice and communication feminization training and the other (tcomm14) was approximately one year into her gender affirmation process and had not received gender affirming services with the exception of hormone replacement therapy.

Byrne (2007) offers a framework that may at least partially explain the results of the present study and appears to fit the presentation of the participants just described. Byrne presents a framework that describes identity development and gender expression as a function of three

stages of affirmation or transition: pre-transition, transition, and post-transition. The transition stage was further divided to include early transition, which was the period of time between the commencement of social affirmation and the completion of GAS. In the transition stage, the affirmed identity is just developing and correct gender attribution often is very important. Tcomm 14 was in the early transition phase and given that she had not yet had voice and communication training or GAS, it makes sense that she would display less confidence in her gender presentation as reflected by her ratings. Furthermore, GAS previously has been associated with both self-perception of voice and QoL (Byrne, 2007; Murad et al., 2010). Other aspects of the framework, specifically post-transition, help to explain the presentations of tcomm1 and tcomm3.

Post-transition is marked by the development of a true sense of self and individuals are more comfortable and accepting of themselves. Being correctly gendered often becomes less important and individuals become more focused on what Byrne terms "getting on with life" (GOWL), in which life goals and identity move beyond being defined by the gender affirmation process. For individuals in this stage, it would make sense for acoustic or gestural characteristics to have less association with perceptions of communication and QoL. Both tcomm1 and tcomm3 were in the post-transition stage; however, only tcomm1 appeared to have moved to GOWL as evidenced by their disparate ratings. Byrne posited that personal factors such as hanging on to the gender affirmation process and having unrealistic expectations could act as barriers to GOWL. This would explain comments made during the data collection session by Tcomm3 suggesting that these types of personal factors were indeed barriers for her. For example, she expressed a lack of confidence in her voice, concern regarding her ability to be correctly attributed despite also

reporting she was not often misattributed, and she inquired about the need for further SLP intervention. This participant rated her satisfaction with both her voice and overall communication as "not at all satisfied".

Limitations and Future Directions

The present study was limited by the lack of representation of feminine communication presentations as determined by available cisgender acoustic norms. For example, only two participants spoke with an f_o that was within the normative range for cisgender women (Baken & Orlikoff, 2000) and only two others were within the gender neutral range of 150-185 Hz (Mordaunt, 2006). Similarly, this study only included individuals who identified as transgender women. Different or stronger relationships may have been evident with greater representation across the spectrum of gender presentation and diversity in gender positioning. Future research should explore whether the same relationships exist across a more diverse sample of participants.

Similarly, participants in the present study all were Canadian and predominantly of Western European (mostly British and French) or Eastern European (i.e., Polish and Ukrainian) ancestry. Given the lack of cultural diversity in the sample, generalization of the present study results may be limited to individuals with similar backgrounds. Replication across varying cultural contexts would be valuable.

Future research also should explore relationships between communication variables and subjective ratings as a function of stage in the gender affirmation process. Communication factors may be more important to and have a more pronounced influence on outcomes for training such as communication satisfaction and QoL for individuals who are in earlier stages than for individuals in later stages. Likewise, individuals in earlier stages of the gender affirmation process may wish to focus training goals on modification of communication

characteristics whereas in later stages, the focus may shift to other alterable factors such as developing strategies to minimize occurrence of and negative responses to misattribution (Azul & Hancock, 2018).

Finally, although this study provides preliminary evidence for relationships between gesture use and certain communicator-centered subjective ratings, further research is needed to explore how gestures are related to gender attribution given its potential influence on subjective ratings, especially in earlier stages of the gender affirmation process. Such information could be used to inform goal-setting (e.g., modifying body language to better align with desired gender presentation and/or to facilitate eliciting desired gender attribution) and clinician feedback (e.g., reinforcement of spontaneous gestural behaviors that are congruent with desired gender presentation).

Conclusions

The aims of the present study were met via the identification of relationships between hand positioning and ST range with ratings of gestural femininity, communication satisfaction, and overall QoL. The study findings suggest that acoustic measures may be less related to client-centered outcomes than previously thought and draw attention to the role of nonverbal communication. This area of research remains underexplored and warrants further investigation with individuals across the continuum of gender identity, presentation, and stage in the gender affirmation process.

CHAPTER 5: LET ME SEE YOUR BODY TALK: AN EXPLORATION OF GENDER DIFFERENCES IN UPPER BODY NONVERBAL COMMUNICATION BEHAVIORS

Abstract

Purpose: The purpose of the present study was to identify group differences in nonverbal communication behaviors based on attributed gender and gender positioning. This was the first study to investigate gender differences in gestural communication that includes transgender participants.

Method: Motion capture technology was used to record the upper body positioning and movements of transgender women (n=20) and cisgender (n=20) communicators during a cartoon retell task. Four gestural variables were measured within a 30-45 second sample of the cartoon description and included total head movement along the mid-sagittal plane; total movement of the lead hand; percent of time the lead hand was in the palm-up position; and elbow position relative to the shoulders. The motion capture data were used to create point-light displays (standardized for size) of the communicators that subsequently were rated by a group of observers (n=20). Observers were asked to indicate the gender of the communicator by selecting from man, woman, and can't decide. Group differences in gesture use were explored using a series of one-way analysis of variance (ANOVA) tests.

Results: Gender attribution ratings were not reliable; therefore, differences between attributed gender groups could not be investigated. However, groups based on gender positioning were found to significantly differ in the use of vertical head movements. Cisgender women moved significantly more on average (492.80°) than transgender women (299.08°).

Conclusions: The results provide preliminary support for targeting nonverbal communication behaviors in communication modification training. The role of these behaviors in cueing gender attribution requires further investigation.

Introduction

Understanding and awareness of gender diversity has gained momentum in recent years and with it has come a quickly evolving lexicon. It is important for those who work and study in the area of gender affirmative care to remain informed about terminology changes in order to be respectful and inclusive in their language use, particularly as the field continues to emerge. Misuse of terminology can lead to misunderstanding and therefore, impede information sharing and also can offend or invalidate the lived-experience of individuals or groups of people. In an effort to enhance clarity, we will begin by defining key terms used in the present study. *Gender performativity* refers to the ways one experiences and produces or expresses their gender as influenced by socially and culturally constructed expectations and norms (Butler, 1990); *gender positioning* refers to one's subjective sense of their own gender; *gender presentation* refers to the ways in which one expresses their gender to others; and *gender attribution* refers to how one's gender is perceived by others (Azul & Hancock, 2018). The focus of the present study was on a specific aspect of gender presentation, nonverbal communication, and how it varies as a function of gender positioning and attribution.

Nonverbal communication behaviors such as gestures represent one of myriad means of realizing gender performativity and have been uniquely described as "…sex in movement" (Zeig, 1985, p.22). Gestures are believed to be gendered (Briton & Hall, 1995; Zeig, 1985), typically according to the binary framework that dominates Western culture, and can be learned by those who are not expected to, or in some cases permitted to, use them (Zeig, 1985).

Information about gesture origin and use has implications for social affirmation and the associated training provided by speech-language pathologists for transgender and gender diverse (TG/GD) people who wish to modify their communication behaviors to better represent their desired gender presentation.

Multiple aspects of gender presentation can be targeted in communication modification (i.e., feminization/masculinization) training including behaviors related to voice, speech, and nonverbal communication. Whereas there is growing evidence that supports voice and speech-based targets, there are limited studies investigating nonverbal communication (e.g., nodding and other head movements, hand gestures, body postures) that have included TG/GD participants (Davies et al., 2015). It is, therefore, important to determine whether or not there are, in fact, differences based on gender positioning in these purportedly socially constructed behaviors and how this constructivism is manifested in TG/GD people. It also is arguably even more valuable to gain an understanding of the ways these behaviors influence gender attribution given the negative consequences of misattribution and/or being identified as a TG/GD person on social-emotional wellbeing, health and safety (Bauer & Scheim, 2015; James et al., 2016b). Improving our understanding of the contribution of gestures to gender attribution also is important because of the role of correct gender attribution as a client-identified training goal. Moreover, nonverbal behaviors previously have been associated with self-rated femininity, communication satisfaction and quality of life (QoL) for transgender women (Hardy, Rieger, Wells, & Boliek, in preparation; Pasricha, Dacakis, & Oates, 2008). Such information can inform clinical practice and the selection of gestural training targets and is the focus of the present study. With the

current dearth of research investigating nonverbal communication in groups of TG/GD participants, clinical practice has been informed, in large part, by social stereotypes and the cisgender literature.

Gender Positioning Differences in Nonverbal Behaviors

Gender differences have been reported in a variety of nonverbal behaviors including amount of interpersonal space used during communicative interactions, body posture, eye gaze, facial expression (especially smiling), use of touch (Hall, 2006; Kennedy & Camden, 1984), nodding (Ashenfelter et al., 2009; Boker et al., 2011; Helweg-Larsen et al., 2004), and use of hand gestures (Bente et al., 1998; Hostetter & Hopkins, 2002; Jones et al., 1999; Nicoladis et al., 2007). Differences in gesture use and head movements were of particular interest for the present study.

Gestures and head movements. Studies have reported that gender differences exist with respect to frequency and type of gestures used. For example, two North American studies reported trends that women used more hand gestures during a story retell task than men (Hostetter & Hopkins, 2002; Nicoladis et al., 2007). Another study, based in Australia, found that women used fewer communicative gestures than men in a conversational context (Jones et al.,1999). Although these studies did not find the same differences, they both support the claim that there are gender differences in gesture use.

Research investigating gender differences in types of gestures used is limited. One study suggested women use more palm-up gestures whereas men use more pointing behaviors (Peterson, 1976). Clinical resources for communication modification training provide additional information about potential differences based, in part, on the general audience work of linguistics, communication, and body language experts as well as the experience of a transgender

woman who developed early resources to guide gender affirmation. These clinical and community-based resources have reported that women keep the elbows in closer proximity to the trunk than men and suggest similar posturing may be targeted with TG/GD clients when gesturing if increased femininity of gestures is desired (Hirsch & Van Borsel, 2012). Although gender differences in this body position have not yet been investigated systematically, the effect of such a posture may be the appearance of a smaller (i.e., more feminine) frame (Hirsch & Boonin, 2012). Conversely, more expansive postures (i.e., elbows away from the trunk) have been associated with greater perceived power or dominance (Carney, Cuddy, & Yap, 2010), considered by some to be a masculine trait (Henley, 1977, 1995).

As with hand gestures, studies have shown that women nod or move their heads in the mid-sagittal plane (i.e., vertical direction) more frequently and with greater amplitude than men (Ashenfelter et al., 2009; Boker et al., 2011; Helweg-Larsen et al., 2004; Jones et al., 1999). These differences existed in both listening and speaking roles.

Overall, there is limited research systematically exploring nonverbal communication as gendered behavior. Similarly, there is only a small body of research investigating the influence of these behaviors as cues to gender attribution.

Nonverbal Behaviors as Cues to Gender Attribution, Masculinity, and Femininity: Evidence from Biomotion Research

Although appearance (e.g., clothing, hair, facial characteristics) is a well-established aspect of gender performativity (Butler, 1990), certain types of motion may cue gender in addition to or in the absence of this aesthetic information. Biomotion research has shown that observers are able to accurately attribute a person's gender based on point-light displays (PLDs) captured during walking (Kozlowski & Cutting, 1977; Mather & Murdoch, 1994; Pollick et al., 2005; Troje,

2002b), orchestral conducting (Wöllner & Deconinck, 2013), and joke-telling activities (Hill & Johnston, 2001). For example, Hill and Johnston (2001) found rigid head movements such as nodding, shaking, and tilting cued speaker gender when presented via animated standardized heads. Similarly, Boker and colleagues (2011) reported that masculinity-femininity ratings of avatars engaged in conversation were based on a combination of appearance and motion (as compared to sex judgments, which were based solely on appearance). Unfortunately, evidence for the role of nonverbal behaviors in effectively conveying gender-related information remains equivocal. Hardy and colleagues (under review, *Journal of Speech, Language and Hearing Research*) reported no significant nonverbal predictors of masculinity-femininity ratings of cisgender and transgender communicators when these variables were entered into a regression model along with voice-related variables. Vertical head movements, however, did significantly predict masculinity-femininity ratings when considered in isolation. These results suggest voice-related aspects of communication are more salient in cueing gender attribution than nonverbal features. Nonverbal cues did, however, still appear to influence masculinity-femininity ratings in that study based on the finding that observers rated communicators significantly more masculine when gestural information was available (i.e., PLD) than when ratings were based on the voice alone.

Taken together, evidence from nonverbal communication, transgender, and biomotion literature suggests that we need to know more about the role of nonverbal behaviors in gender performativity and attribution and that gender differences (actual and perceived) in use of hand/arm gestures and head movements are worthy of further exploration.

Purpose

The purpose of the present study was to explore gender differences in nonverbal communication behaviors in order to inform communication modification training for TG/GD people. Specific questions related to this purpose included:

1. Are there differences in upper body nonverbal communication behaviors (i.e., hand movements, hand position, elbow position) between groups based on
    a. Attributed gender?
    b. Gender positioning?
2. Are there differences in head movements between groups based on
    a. Attributed gender?
    b. Gender positioning?

Methods

This study was one of four in a line of research investigating relationships between communication behaviors and subjective ratings of gender presentation and attribution. Detailed descriptions of methods can be found in Hardy et al. (2018) and Hardy et al. (under review, *Journal of Speech, Language and Hearing Research*). Key information related to the present study is summarized here for the reader's reference. Ethical approval for this study was obtained from the University of Alberta Health Research Ethics Board (HREB) (Pro00060133) prior to recruiting participants. Operational approvals also were obtained from the Northern Alberta Clinical Trials and Research Centre (NACTRC) and Covenant Health Research Centre.

Participants

Two groups of participants were recruited from Edmonton and surrounding areas for the present study. The first group served as communicators (n = 40) and provided the gestural data. The second group served as raters and provided the gender attribution data. The communicator group included transgender women (n = 20) and age-matched cisgender women and men (n = 10 each). The mean age of the communicators was 40.93 years (SD = 14.48). Each participant was required to have full use of their upper body and not have any neurological conditions that may have caused them to gesture differently than typical speakers. Transgender participants self-identified as transgender women and had been living as their true, feminine selves the majority of the time (i.e., at least 80% of the time) for at least six months. The majority of these participants (60%) had not participated in communication modification training with a speech-language pathologist, although many were planning to access those services in the future. The present study included only transgender women as they represent the bulk of the communication modification clinical caseload and population-specific studies are important for informing clinical practice.

Raters self-identified as women (n = 10) and men (n = 10) and were 29.26 (SD = 9.14) years old. They were required to have adequate vision and cognition, per self-report, to complete the rating tasks. Raters were blind to the nature of the experimental objectives and the inclusion of transgender participants. They were debriefed at the end of the experimental tasks.

Recruitment methods for all participants included convenience sampling, word-of-mouth advertising, and a media release. Transgender participants also were recruited from the voice program at the Glenrose Rehabilitation Hospital (former, current, and prospective clients) via invitation letters and posters.

Gestural Data Acquisition

Motion capture recordings were obtained of the communicators' gestures while they retold the story of a short cartoon (Davis, 1969). A set of 15 custom passive motion markers (18 mm on a two mm base) was attached to standardized locations on each participant's head and major upper body joints. These locations included the center of the forehead, temples, spinous process of the seventh cervical vertebra, sternum, shoulders, elbows, wrists, and the center of the hands (see Figure 5.1). A Motion Analysis Corporation 8 camera optical motion capture system recorded the marker movements at a frame rate of 120 Hz.

Figure 5.1. Passive Motion Marker Placement

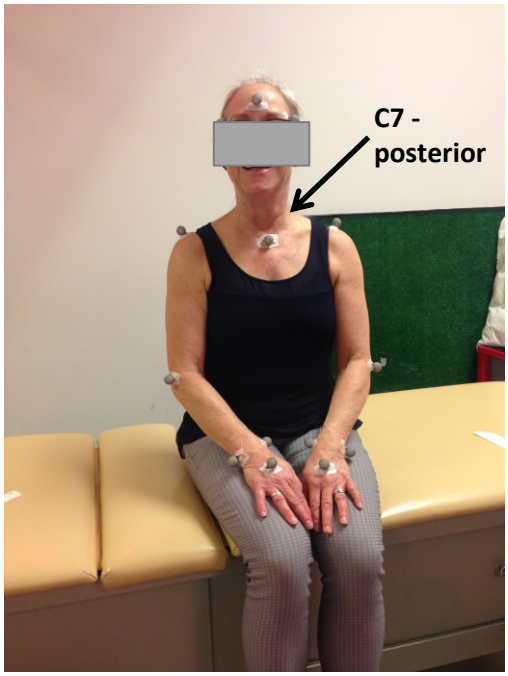


Figure 5.1. Placement of passive motion marker set (15 markers) on a member of the research team. The C7 marker (not visible) was placed on the dorsal aspect of the base of the neck on the spinous process of the seventh cervical vertebra.

Participants were seated on a raised stool at a set distance from and facing the first author and retold the story of the cartoon they had just viewed. A raised stool was chosen to allow for free range-of-movement of the arms, elbows, and hands, and to prevent posturing that would have limited gesturing (e.g., leaning forward with the elbows on the legs). A cartoon description task was chosen because of its narrative nature and ability to elicit gestures (D. McNeill, 1992). Finally, the second member of the dyad was held constant (i.e., the first author) to control for influences secondary to the interaction. For example, people use a greater number and extent of head movements when conversing with a woman (Ashenfelter et al., 2009; Boker et al., 2011).

Gesture measurement. Four variables were selected to represent head/hand movements and hand/elbow positions for the present study. These included total head movement in degrees along the mid-sagittal plane (MS Head); total movement in millimeters (mm) of the lead (i.e., dominant) hand irrespective of direction (Hand Mvt); percent of time the lead hand (i.e., the hand that moved the most) was in the palm-up versus palm-down position (Palmup%); and mean percent elbow:shoulder width (Elbow%).

A customized Matlab script was written to extract these measurements from the X, Y, and Z coordinate data contained in the motion capture recordings (i.e., .trc files). Definitions of each variable can be found in Hardy et al., (under review, *Journal of Speech, Language and Hearing Research*). Measurements were obtained from a 30-45 second sample of the cartoon description extracted to control for inter-communicator differences in length and complexity of the description. Gesture rate has been associated with linguistic complexity; therefore, previous authors have recommended linguistic complexity be controlled when making comparisons (Nicoladis et al., 2007). The sample was selected based on the verbal description that was recorded at the same time as the gestures. Detailed descriptions of the editing protocol can be

found in Hardy, Rieger, Wells and Boliek (2018) and Hardy and colleagues (under review, *Journal of Speech, Language and Hearing Research*). The same 30-45 second sample of motion data was used to create the stimuli for the rating phase of the study.

Rating Data Acquisition

A customized software program ("Gender Finder," 2017) was developed to create and present the stimuli for the rating phase and collect the rating data. The full motion capture recordings of the cartoon descriptions (in .c3d format) first were converted to PLDs in video format that were standardized for size across participants (see Figure 5.2). The resultant .mp4 files were edited to the same 30-45 second sample that was used to obtain the gestural measures.

Figure 5.2. Example of Point-Light Display

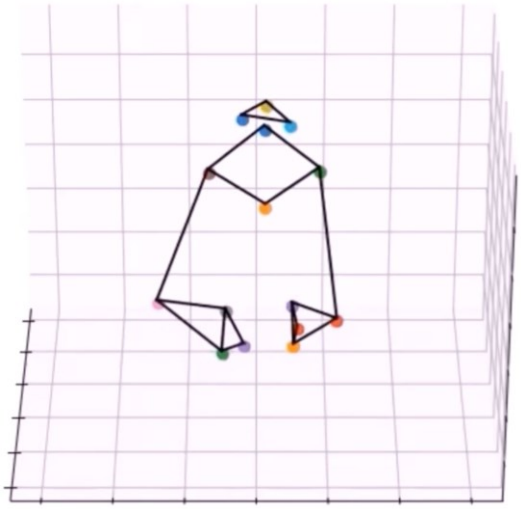


Figure 5.2. Point-light display of a pilot participant while describing the cartoon. The dots correspond to the passive motion markers.

All rating data were obtained in a quiet room at the University of Alberta or Glenrose Rehabilitation Hospital using a laptop computer loaded with the Gender Finder program. After consenting to participate, providing demographic information, and completing standardized training activities, participants commenced the research rating tasks. Gender Finder randomly

presented the PLD stimuli of all 40 communicators immediately followed by files repeated for reliability measurements, which also were presented in random order. After a stimulus item finished playing, Gender Finder presented the rating platform on screen. Raters were asked to attribute gender of the communicator by selecting from a drop-down menu consisting of three options: man, woman, and can't decide. The ratings were automatically entered into a spreadsheet. Gender Finder was extensively tested prior to commencing data collection to ensure stimulus presentation was randomized appropriately and rating data were stored accurately. The rating task took approximately 25-35 minutes to complete. Communicators were assigned an attributed gender if they were consistently attributed a particular gender by raters (i.e., at least 80% of the time, as in Gelfer and Bennett, 2013). If a communicator was not rated consistently as either man, woman, or 'can't decide', they were assigned 'can't decide' as their gender was considered to be ambiguous to raters.

Statistical Analyses

Group differences were identified using a series of one-way analysis of variance (ANOVA) tests. Hand Mvt, Palmup%, and Elbow% served as the dependent variables (DV) for the first question and gender groups served as the independent variables (IV). Each IV had three levels. Attributed gender included man, woman, and can't decide and gender positioning groups included cisgender man, cisgender woman, and transgender woman. MS Head served as the DV for the second question and the IVs remained the same as the first question. A Bonferroni correction was applied to analyses involving upper body gestures given that multiple analyses were undertaken ($\alpha = .016$). No correction was applied for analyses involving MS Head ($\alpha = .05$).

Results

Reliability of Rating Data

Raters were not able to reliably attribute gender based on gestural cues alone. Intra-rater reliability, as measured by percent agreement between first and second ratings of the four repeated files, was only 60%. Inter-rater reliability was calculated using Fleiss' Kappa and was $\kappa = .11$. Given the lack of reliability of the rating data, differences between attributed gender groups could not be investigated in the present study.

Differences in Nonverbal Communication Behaviors Based on Gender Positioning

The results of the one-way ANOVAs revealed no significant differences between gender groups for Hand Mvt $F(2,37) = .014$, $p = .987$, Palmup% $F(2,37) = .959$, $p = .393$, or Elbow% $F(2,37) = .347$, $p = .709$. There was, however, a significant gender effect for MS Head $F(2,36) = 5.06$, $p = .012$. See Table 5.1 for descriptive statistics of the four gestural variables. Effect size calculations revealed 21.9% of the variability in head movements could be accounted for by gender identity (partial eta squared = .219). Post-hoc testing (Tukey HSD) revealed cisgender women used significantly more vertical head movements than transgender women ($p = .009$). Cisgender women also used more vertical head movements than cisgender men; however, this difference was not significant ($p = .079$). See Figure 5.3 for a graphic display of the results.

Figure 5.3. Gender Differences in Total Mid-Sagittal Head Movement

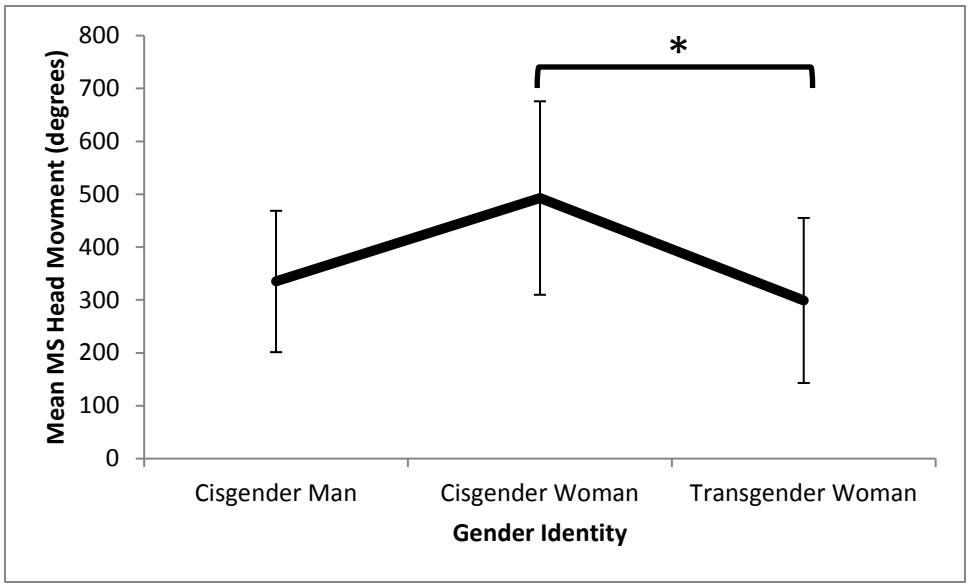


Figure 5.3. Mean total head movement along the mid-sagittal plane presented as a function of self-identified gender group. Significant group differences (α = .05) are marked with an *. Error bars represent +/- one standard deviation. Note: An extreme outlier (Tcomm13) was excluded from the analysis resulting in a sample size of n = 19 for the Transgender Woman group.

Table 5.1 *Descriptive Measures for Gestural Variables Presented by Gender Groups*

Gestural Measure	Transgender Women (n = 19) Mean (SD)	Cisgender Women (n = 10) Mean (SD)	Cisgender Men (n = 10) Mean (SD)
MS Head (Degrees)	299.08 (156.06)	492.80 (182.94)	335.01 (133.62)
Hand Movement (mm)	5123.59 (5370.18)	8039.87 (4549.85)	6297.84 (6373.41)
Palmup%	17.72 (26.01)	12.55 (14.07)	21.75 (29.94)
Elbow%	144.45 (26.16)	143.84 (13.40)	143.06 (18.09)

Note. An extreme outlier was excluded from MS Head in the Transgender Women group.

Discussion

The aims of the present study were to identify group differences in nonverbal communication behaviors based on attributed gender and gender positioning. Gender attribution data were not reliable; therefore, those group differences could not be explored. Gender positioning groups were found to significantly differ in use of vertical head movements only. These movements represented any change in position along the midsagittal plane and were not necessarily associated with a purposeful gesture such as nodding, although such gestures also would have been captured in the measurement.

The results of the present study replicated findings of previous studies reporting cisgender women use frequent vertical head movements while communicating (Ashenfelter et al., 2009; Boker et al., 2011; Helweg-Larsen et al., 2004; Jones et al., 1999). Cisgender women in the present study moved their heads, on average, 492.80° (SD = 182.94°) over the course of a 30-45 second narrative task. This value was 157.79° and 193.72° greater than that for cisgender men and transgender women, respectively. These findings provide the first empirical support for modifying nonverbal behaviors in communication modification training programs for TG/GD people.

The finding that the only significant group difference was between cisgender and transgender women was unexpected given that many transgender women believe their body language to be feminine (Hardy, Rieger, Wells, & Boliek, in preparation; Pasricha et al., 2008). The group difference and discrepancy between perceived and measured behaviors may be explained by gender performativity and socialization with gender normative behaviors. The transgender women participants in the present study had been socialized in masculine gender roles during their formative years (Rehman, Lazer, Benet, Schaefer, & Melman, 1999) and as such, societal

expectations would have dictated the development of "typical" masculine mannerisms (Gagné & Tewksbury, 1999). Indeed, some individuals may even have developed hyper-masculine behaviors in an attempt to mask their natural tendencies and true gender positioning (Byrne, 2007; Gagné & Tewksbury, 1999). After years of expressing a masculine gender presentation, it is common for masculine behaviors to persist into the time of gender affirmation (Byrne, 2007). This persistence offers a possible explanation for the transgender women using less head movement than the cisgender women and a nearly identical amount as cisgender men in the present study, especially since more than half (i.e., 60%) had not yet participated in voice and communication feminization training at the time of data collection.

The lack of significant gender differences in upper body gestures also was unexpected given the existing evidence for differences in gesture use amongst cisgender communicators (Hostetter & Hopkins, 2002; Jones et al., 1999; Nicoladis et al., 2007; Peterson, 1976; Yang, 2010). Visual inspection of the present results showed that cisgender women moved the lead hand more and used palm-up hand position less than cisgender men. Transgender women used even less hand movement than cisgender men but positioned the hand palm-up for a percentage of time somewhere between that of cisgender women and men. These observed qualitative differences may have failed to reach significance due to the large degree of within-group variability in gesture use as evidenced by the large standard deviations. Elbow position for all three groups virtually was the same; therefore, the lack of significant difference likely represents a true result in that case. The gestural characteristics of the transgender women participants add additional credence to the hypothesized role of socialization and gender performativity but also suggest that some behaviors may indeed be at least somewhat feminine for these communicators.

The gestures selected for the present study represent only a subset of those possible and were limited to gross movements due to the exploratory nature of the study. Other gestures or movements may have revealed additional gender group differences. For example, fine movements, such as those associated with facial expression, are thought to differ between cisgender communicators (Hirsch, 2006) and have been shown to cue gender in averaged avatars created from cisgender men and women (Hill & Johnston, 2001). These types of movements offer additional avenues of exploration for future studies.

A Hawthorne effect as proposed in Hardy and colleagues (under review, *Journal of Speech, Language and Hearing Research*), may provide another explanation for the lack of significant gender differences in upper body gestures in the present study. In that study, we investigated communication-based predictors (i.e., characteristics of voice, speech, and gestures) of masculinity-femininity ratings made of the same group of communicators by the same group of raters. None of the measured nonverbal behaviors significantly predicted the subjective ratings in that study. We postulated that some communicators acted unnaturally (compared to their typical behavior) secondary to the recording condition. For example, several communicators were observed to gesticulate freely and frequently during informal conversation but not during the cartoon description. Being recorded likely made participants more aware and self-conscious of their behaviors, resulting in fewer gestures than would be typical for their communication patterns. Explicit instructions to "gesture as you normally would" were not provided so as to minimize this awareness; however, some participants may have believed they needed to sit still for the recording in order to be a "good research participant". For example, one cisgender participant stated she did not move her hands at all because she thought she was not supposed to move them. Moreover, the markers themselves likely brought awareness to limb and hand

movements that would not otherwise have been present. Ten markers were attached to the arms and hands (five on each side) using adhesive electrode washers and surgical tape. There was no discomfort associated with the marker attachments; however, it was not a routine experience for any of the participants and several of the markers were visible to them during the cartoon description. For a few participants, this awareness was heightened by loss of adhesion of the surgical tape during data collection secondary to lotion application prior to attending the study or perspiration. Future researchers may wish to explore possibilities for methodologies involving naturalistic conversation or consider making instructions more explicit that movement is allowed during recording while still minimizing attention to specific movements (e.g., "Don't feel like you can't move. The cameras are all around the room so they'll still be able to see you.").

The failure to fully replicate previous findings suggests we do not yet completely understand gender differences in nonverbal communication and justifies additional research. Moreover, results may have been different between attributed gender groups than groups based on gender positioning because, as those who have experienced being misattributed can attest, gender positioning and attribution are not always congruent. Given the continued pervasiveness of societal expectations for presenting or performing one's gender and the documented beliefs about what constitutes masculine or feminine nonverbal communication (Briton & Hall, 1995), one would expect that attributed women and men would differ in behaviors such as use of hand gestures.

Although gender differences are interesting and informative, differences between attributed gender groups arguably would have been more relevant for communication modification training practices because of the information they would provide regarding the influence of gestures in gender attribution outcomes.

The lack of reliability of these ratings in the present study was, in itself, an important result. It suggests gestures may not be strong enough cues to gender when considered in isolation, without other aspects of presentation such as appearance or voice. It is possible there is not enough information in gestural communication for raters to make consistent attributions. Whereas it generally is accepted that PLDs contain enough information to convey gender and that observers can use this information to correctly recognize gender (Pollick et al., 2005), the results of the present study suggest that may not be the case or that accuracy and/or reliability may be activity or presentation specific.

A meta-analysis investigating gender recognition of point-light walkers revealed that observers were 61-71% accurate on average, depending on the view of the walker (Pollick et al., 2005). The highest accuracy reported was 86% (Runeson & Frykholm, 1983). Gender recognition accuracy based on other movements (e.g., knocking, waving, lifting, musical conducting), however, has been reported as lower than that for walking (Pollick et al., 2002; Wöllner & Deconinck, 2013). Pollick and colleagues (Pollick et al., 2005, 2002) suggested human observers may not use all available information when making ratings based on motion and estimated human efficiency for gender recognition at 0.27% for arm movements and 26-47% for walking when compared to an ideal rater (i.e., computer-based neural network). Percent agreement calculated for gender attribution in the present study was similar to previous estimates of gender recognition accuracy just described.

Despite human limitations in making perceptual ratings, there is evidence to suggest motion information alone is sufficient to cue gender. Hill and Johnston (2001) reported both rigid (e.g., vertical head movements) and non-rigid head movements (e.g., facial expression) contained salient cues to a communicator's gender and raters correctly attributed gender 61.9% of the time.

In their study, motion information was presented using motion capture-based animation techniques, which offer similar control for appearance and body morphology cues as point-light displays through stimulus averaging or standardization. It would be interesting to compare gender attribution ratings based on these two visual presentation methodologies to determine the one that yields the most reliable ratings.

Although it is possible that observers are not able to reliably attribute gender based on gestures, the reader should be cautious with respect to drawing firm conclusions based solely on the results of the present study, especially given the results of Hill and Johnston (2001) just described. There may have been reasons other than lack of salience of motion-based information contributing to the low reliability. One such possibility is the aforementioned Hawthorne effect and its potential limitation on movements. This limitation in combination with only having PLDs from which to judge may have affected the attribution ratings. If the communicators had moved more (e.g., used more hand gestures), raters would have had more movement to judge and the point-light displays may have been more "potent" and thus, reliable.

Another possibility is "bias-checking" on the part of the raters as proposed in Hardy and colleagues (under review, *Journal of Speech, Language and Hearing Research*). Bias checking was described as becoming aware of personal biases and stereotypical beliefs and subsequently avoiding using those preconceptions to guide ratings/attributions. Whereas bias checking may be desirable from inclusivity and social justice perspectives, those behaviors may not yet reflect reality, especially when considering the lived experiences of TG/GD people. Expectations about gendered behaviors often go unchecked during natural social interactions and observations. Indeed, studies have shown people are able to accurately categorize members of perceptually ambiguous social groups (e.g., gay men) at rates higher than chance based on stereotypic

expectancies related to movement, speech, appearance, and so on (Tskhay & Rule, 2013). The self-monitoring behaviors observed in the present study may have, consequently, masked the existence of true socially constructed groups and the gestural cues actually used by conversation partners to attribute the gender of those with whom they interact. As suggested in Hardy and colleagues (under review, *Journal of Speech, Language and Hearing Research*), instructing raters to make instinctual ratings and providing assurance about the acceptability (for the purpose of the study) of using personal beliefs to guide ratings may have improved reliability. Finally, another option is to recruit participants based on known and varied attribution experiences (e.g., those that are consistently correctly attributed and misattributed) rather than grouping based on ratings. This approach also would facilitate obtaining equal group sample sizes.

Conclusion

The present study was the first to systematically investigate gender differences in nonverbal communication behaviors including transgender participants. Overall, the results suggest that individuals who were socialized in the girl/woman gender role use more vertical head movements when communicating than those raised in boy/man role. These findings provide preliminary support for targeting nonverbal communication behaviors in communication modification training with TG/GD people with one focus being on head movements. The role of these behaviors in cueing gender attribution remains unclear and is worthy of further investigation.

CHAPTER 6: CONCLUDING PERSPECTIVE

The gender affirmation process for transgender/gender diverse (TG/GD) people can involve a number of different socially (e.g., name change, clothing choices) or medically based changes (e.g., surgery) (E. Coleman et al., 2011). The nature of these changes and associated services accessed are determined by the needs and desires of the individual in terms of their declared gender positioning and their wishes for how to present that positioning to others, as well as how they want others to attribute their gender. For many, modifying communication behaviors is important in order to better align gender presentation and attribution with the way in which they identify, according to their own beliefs as well as social and cultural expectations.

Incongruence between voice and communication characteristics and gender positioning can be a cause of gender dysphoria and lead to misgendering or being outed as a TG/GD person, which can have a negative influence on quality of life (QoL), life participation, and even personal safety (Bauer & Scheim, 2015; Hancock, 2017; Hardy et al., 2013; James et al., 2016b). For these reasons, it is important to identify the aspects of voice and communication that are most salient in conveying gender and masculinity-femininity. Once identified, these characteristics can serve as individualized training targets in communication modification training programs. Moreover, the information will better equip speech-language pathologists (SLPs) to guide TG/GD clients in achieving a gender presentation that is more congruent with their stated gender positioning in a way that sounds and/or feels authentic or natural to them and others. Limited resources for services and client-centered care practices provide additional impetus for streamlining practice and avoiding spending precious clinical time focused on aspects of communication that are not salient cues to gender or masculinity-femininity.

Summary of Contributions

The aim of my doctoral research was to identify a set of communication-based predictors (i.e., acoustic and nonverbal communication measures) of subjective ratings related to gender attribution. More specifically, I wanted to advance our understanding of communication features that serve as cues to others about gender, masculinity-femininity, and vocal naturalness. I also explored relationships between communication characteristics and self-rated outcomes of femininity (of voice, gestures, and overall communication), communication satisfaction (voice-specific and overall), and quality of life (QoL) (general and voice-related) for transfeminine communicators (a.k.a., transgender women). Finally, I investigated differences in a set of nonverbal communication behaviors between groups based on gender positioning and gender attribution. I addressed my research aims via a series of four studies. The main findings of these studies are summarized below:

1. Speaking fundamental frequency (f_o) was the sole voice-based predictor of gender attribution. The higher f_o, the more likely a communicator was to be attributed a gender other than man.
2. F_o, average formant frequency, and sound pressure level (SPL) significantly predicted masculinity-femininity ratings made by raters in both audio only and audiovisual (AV) modes. F_o and average formant frequency were positively related to ratings whereas SPL was negatively related to ratings. The higher the f_o and average formant frequency, the more feminine the ratings. Conversely, the higher the SPL, the more masculine the ratings. No gestural variables significantly predicted observer ratings when considered with acoustic variables.

3. F_o, average formant frequency, and rate of speech significantly predicted voice naturalness ratings. F_o was negatively related to the subjective ratings whereas average formant frequency and rate of speech were positively related to the ratings. The lower the f_o, the higher average formant frequency, and the faster the speaking rate, the more natural the ratings.
4. Use of palm-up hand gestures was negatively correlated with self-rated gestural femininity and satisfaction with overall communication for transgender women communicators. Positioning the hands with palms facing downward was associated with higher gestural femininity ratings and greater satisfaction with overall communication. Self-rated overall QoL was positively related to mean semitone range across utterances included in the cartoon description sample. The higher an individual rated their QoL, the greater the variability in their intonation patterns. There were no significant relationships identified between any of the acoustic variables and voice femininity, satisfaction with voice, or voice-related QoL.
5. Cisgender women moved their heads along the mid-sagittal plane (i.e., up and down) significantly more than transgender women while retelling a story to another person. These movements included, but were not limited to, gestures such as nodding.

Together these interrelated studies have advanced our understanding of the role of communication characteristics as predictors of gender attribution, masculinity-femininity, and naturalness ratings. They also have increased our understanding of the ways in which these characteristics relate to self-attributed masculinity-femininity, QoL, and communication satisfaction. My research findings have made contributions in the areas of voice and speech

acoustics, gestural communication, observer characteristics, gender-based communication characteristics considering non-binary positioning and attributions, and QoL. These contributions are summarized below:

1. The studies included in my doctoral research investigated a unique constellation of communication variables. Previous studies in the voice and communication modification literature have investigated the contribution of certain acoustic variables to perceptual measures related to gender; however, most focused on f_o, vowel formant frequency, and/or intonation. The unique set of acoustic variables included in Studies 1 (Hardy, Rieger, Wells, & Boliek, 2018), 2 (Hardy et al., under review, *Journal of Speech, Language, and Hearing Research*), and 3 (Hardy, Rieger, Wells, & Boliek, in preparation) represented voice and speech more comprehensively than previous studies. For example, these studies included SPL, and rate of speech (RoS), which rarely have been included previous variables sets. Moreover, the communication variables included in my doctoral research were not limited to acoustic measures: Study 2, 3, and 4 (Hardy et al., in preparation) also contained measures of nonverbal forms of communication including gesture.
2. The results of Study 1 and 2 confirmed the importance of f_o as a cue to gender attribution and masculinity-femininity. Relationships between f_o and those subjective measures are well-documented for cisgender and transgender communicators (Leung et al., 2018) and both studies found the same relationships existed for the present group of participants. The results of those studies also supported previous findings suggesting vocal tract resonance characteristics and SPL contribute to perceived masculinity-femininity and replicated previously-reported relationships (Leung et al., 2018). Moreover, these studies

contributed to the limited population-specific evidence base. Whereas vowel formant frequencies have been fairly widely studied in terms of their contribution to gender attribution and/or perceived masculinity-femininity, most studies included only cisgender participants (Leung et al., 2018). Studies exploring the role of SPL are very limited in number and only one previously included TG/GD participants (Holmberg et al., 2010). The results of Study 1 and Study 2 support the continued focus on f_o and vocal tract resonance as training targets in voice and communication modification training and provide preliminary evidence for increased focus on SPL. Previous research has shown that SPL goals are not common in this type of training (Hancock & Garabedian, 2013).

3. Study 1 identified acoustic predictors of vocal naturalness. No other researchers have measured this variable amongst TG/GD participants yet anecdotal evidence suggests individuals who pursue voice and communication modification training are seeking a natural-sounding or "authentic" voice. The results of Study 1 partially replicated the findings of Hardy et al. (2016) in identifying vowel formant frequency as a significant predictor of vocal naturalness. The total set of significant predictors was not identical between studies; however, the participant groups also varied in terms of gender positioning and access to gender affirmation services. In Hardy et al., (2016), participants were all transgender women who had not yet availed themselves of voice and communication training whereas my doctoral research included both cisgender and transgender communicators, some of whom had already accessed training. The naturalness-specific results of Study 1 provide further support for targeting f_o and vocal

tract resonance in communication modification training as these aspects of voice appear to not only contribute to listener-perceived masculinity-femininity but also the naturalness of that speaker's voice.

4. Studies 2, 3, and 4 were the first to quantify gestural variables in the TG/GD population and systematically investigate their contribution to gender attribution and masculinity-femininity ratings. Furthermore, Study 4 was the first population-specific study to explore gender differences in nonverbal communication behaviors. The lack of population-specific research in this area previously was identified as an area of need (Davies et al., 2015). Together, these studies provide the first evidence to support targeting nonverbal communication behaviors (i.e., head movements, hand position) in training programs and demonstrate that relationships exist between nonverbal communication and self-rated outcomes. Moreover, in addition to the acoustic results, these studies provide important confirmation for the work clinicians are doing intuitively, based on clinical experience, or based on the relatively limited population-specific research evidence.
5. The results of Study 2 and Study 4 revealed that the raters who participated in my doctoral research were not reliable in their gender attribution and masculinity-femininity ratings when ratings were based on visual information alone (i.e., the point-light display). I proposed that the lack of reliability might have been due to "bias checking" or self-monitoring use of stereotypic expectancies when making ratings. Indeed, several raters reported during debriefing that they had engaged in such self-monitoring. Moreover, studies have shown that people are able to accurately categorize members of socially ambiguous groups (e.g., sexual orientation) and attribute gender based on visual/motion

information (Mather & Murdoch, 1994; Tskhay & Rule, 2013; Wöllner & Deconinck, 2013), suggesting ratings should have been reliable. It also is possible the lack of reliability was due to a deficiency of salient information in the visual stimuli on which to make the ratings, either due to the characteristics of the point-light displays or lack of communicator movement. Some communicators appeared to have experienced a Hawthorne effect and consequently gestured less during the research task than they were observed to do in informal conversation. I suggested the reliability issues be addressed in the future by modifying instructions for raters and communicators to foster ratings that better represent typical, unchecked, social interactions and gesturing behaviors, respectively. Future researchers also may wish to consider using motion-capture derived animations or avatars rather than point-light displays to provide more realistic, but still controlled, visual stimuli.

6. My doctoral research explored relationships between communication variables and both observer and communicator-reported (for trans-identified participants) subjective ratings. Most studies in the voice and communication modification literature have included only listener ratings (Note: referring to listener rather than observer because ratings made in previous studies were based on the voice alone). Relationships with observer ratings are important to advance our understanding of the contribution of communication to gender attribution. Self-reported measures, however, also are important for informing client-centered training outcomes, especially because observer and self-reported ratings are not necessarily related to other important outcome measures such as QoL in the same way (Hancock et al., 2011; E. J. M. McNeill et al., 2008).

7. Finally, gender attribution ratings collected in the present research included a third, gender-neutral, option not typically found in the voice and communication modification literature. The inclusion of this category was important because many transgender clients end up with gender-neutral presentations following training and others position themselves as gender-neutral. Gender does not exist in a binary and, therefore, it is important to understand how communication relates to a spectrum of identities, presentations, and attributions.

Limitations

The present line of research included several limitations that should be acknowledged. Given the data collection procedures (i.e., all studies drawing from the same data set), most of these limitations are applicable to all four studies.

Lack of Diversity in Gender Positioning, Presentation, and Attribution

The transgender participants in the present research only included transfeminine communicators. Although this group is representative of most clinical caseloads, clientele is changing, reflecting greater diversity in identity or positioning, desired presentation, and desired attribution. Results obtained from transfeminine communicators may not be applicable to other GD individuals. Moreover, the communication characteristics and resultant observer ratings of this particular group of participants lacked adequate variability. Most transgender participants were rated to have masculine communication behaviors and consequently were attributed as men. The majority also had not yet participated in voice and communication modification training. This lack of diversity limited what information could be gained with respect to communication characteristics that effectively predict feminine ratings amongst communicators who have androgen-altered anatomy. Similarly, the lack of variability may have impacted the

statistical results, limiting the ability to identify communication-based predictors of gender attribution. This line of research should be repeated with participants who represent a diverse spectrum of identities and attributions. For example, very little research has focused on voice and communication modification for transmasculine people or individuals who are nonbinary.

Sample Size and Unequal Attributed Gender Groups

Although the present sample size of 40 was relatively large compared to other studies in the voice and communication modification literature, it was small when considering the complex statistical tests used to address the research aims. This limitation was especially problematic for the logistic regression analyses identifying predictors of gender attribution, as it is desirable to have very large sample sizes or case:predictor ratios when working with categorical criterion variables (Tabachnick & Fidell, 2013). Case:predictor ratios refer to the number of participants for every predictor variable included in the regression analysis. It is calculated by dividing the sample size by the number of predictor variables. The small sample size resulted in underpowered studies. The population of transfeminine people living in the Edmonton area is small. It is estimated that 0.6% of the overall population identifies as TG/GD (Flores et al., 2016) and not all of those people identify as transfeminine. Restrictions in lab availability further limited the number of people who could potentially participate in data collection: the lab was only available during certain business hours and, as a result, excluded people who worked during regular business hours.

Another related limitation was the unequal sample sizes of the attributed gender groups. The previously described lack of variability in communication-based gender presentation resulted in a disproportionate number of participants being attributed as men. Very few transgender

participants were attributed as women or ambiguous in gender, resulting in fewer cases in each of those two groups. Equal group sizes also are desirable for analyses involving categorical criterion variables (Tabachnick & Fidell, 2013).

With a larger sample size, equal group sizes, and thus, increased statistical power, other communication variables may be identified as significant predictors of the subjective ratings included in the present studies. Potential significant predictors include the variables identified in the respective purposeful selection of covariates analyses in Study 1 and Study 2 that were excluded from the final predictive models (i.e., total mid-sagittal head movement, RoS, and laryngeal acoustic measures). Measures related to intonation (e.g., frequency of upward and downward intonation patterns) offer additional possibilities for potential predictors of the subjective ratings (Leung et al., 2018). Future researchers should consider cross-institutional collaborations as a means of recruiting sufficient numbers of participants for adequate statistical power, as it seems unlikely that it is possible to do so from a single setting/location.

Reliability of Measuring Intonation Shifts

Unfortunately, the method chosen to measure intonation shifts in the present series of studies was not reliable. Given the timeline for completing my doctoral research, it was not possible to develop another procedure so this variable could be included in the overall set of potential predictors; therefore, ST range served as an alternate measure of intonation. There was a statistical benefit of excluding this potentially important variable in that the case:predictor ratio already was smaller than was desired in some studies. Reducing the overall number of variables in the set increased the ratio. Nevertheless, a meta-analysis of acoustic features contributing to listener perception of speaker gender identified intonation shifts as potentially important (Leung et al., 2018); therefore, reliable methods of measuring this variable are needed and should be

reported in detail to facilitate replication across studies. Moreover, it will be important to explore the predictive strength of this variable in relation to other factors such as f_o, vowel formant frequencies, and SPL.

Visual-Only Mode Ratings and Nonverbal Communication

Lack of reliability in visual-only mode ratings and potential restrictions in use of nonverbal communication behaviors was discussed in detail in Studies 2 and 4. These papers described a Hawthorne effect and bias checking as potential causes of these limitations. In summary, I suggested that raters became aware of their stereotypical beliefs with respect to gestural communication behaviors and subsequently actively refrained from using those behaviors to inform their subjective ratings. I also suggested that some of the communicators gestured less than they typically would as a result of the recording context. Consequently, the true salience of gestural communication cues may have been masked. Addressing these limitations will help to better elucidate the role of nonverbal communication behaviors in cueing gender attribution.

Future Directions

My doctoral research included both extensions of previous research and novel areas of study. Certain aspects of the research offer direction for further exploration. For example, there are relevant communicator-rated measures that have not yet been explored in relation to communication features. Examples of these subjective ratings include the frequency with which an individual is addressed as the desired gender category and the degree to which communication reflects the individual's gender. Both ratings have been proposed as important client-centered outcomes (Azul & Hancock, 2018); therefore, it would be valuable to understand how they relate to communication features.

Another area that remains underexplored in the TG/GD population is nonverbal communication. Studies 2 and 4 have taken important first steps in investigating hand and head movements but need to be expanded by addressing the limitations of the Hawthorne effect and bias checking. In addition, those studies measured only gross gestures and positions of the upper body. It would be valuable to explore finer aspects of nonverbal communication such as facial expression or lower body characteristics such as postural leg position, which are thought to contribute to gender attribution or the expression of masculinity or femininity (Burns-Ardolino, 2003; Hill & Johnston, 2001).

Finally, the four studies comprising my doctoral research all were observational in nature. Although they provide important information relevant to clinical practice, additional research is needed to more fully inform voice and communication modification training. By addressing the limitations described earlier, we will begin to narrow down those communication features that contribute to gender attribution across the entire spectrum. Identifying those features represents the next critical step in this line of research. Once the features have been identified, the next steps will be to carry out training efficacy and effectiveness studies to determine whether or not making changes to the identified set of communication features results in corresponding changes to subjective attributions and client-rated outcomes. Such studies also will be necessary to identify ideal training-delivery models and dosage to achieve the greatest clinical outcomes possible in the most efficient manner. Efficient service-provision is becoming increasingly important with growing demand for services and limited clinical resources.

Conclusion

The results of my doctoral research advanced our understanding of the ways in which voice and communication characteristics contribute to the spectrum of gender presentations and attribution. Together, they will add to the growing evidence base informing voice and communication modification training and thus have the potential to positively impact quality of life and life participation outcomes for TG/GD individuals.

REFERENCES

519, T., & Ontario, R. H. (2016). Media reference guide: Discussing trans and gender-diverse people. Retrieved from https://www.rainbowhealthontario.ca/wp-content/uploads/2016/01/Media-Reference-Guide-Trans-and-Gender-Diverse-People1.pdf

Abernethy, B., & Zawi, K. (2007). Pickup of essential kinematics underpins expert perception of movement patterns. *Journal of Motor Behavior*, *39*(5), 353–367.

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

Adler, R. K., Hirsch, S., & Mordaunt, M. (Eds.). (2012). *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide* (Second). San Diego, CA: Plural Publishing Inc.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental health disorders: DSM-5 (5th ed.)* (5th ed.). Washington, DC: American Psychiatric Publishing.

American Speech-Language-Hearing Association. (2016). Scope of Practice in Speech-Language Pathology. American Speech-Language-Hearing Association.

Andrews, M. L. (1999). *Manual of voice treatment: Pediatrics through geriatrics* (2nd ed.). San Diego, CA: Singular Publishing Group, Inc.

Andrews, M. L., & Schmidt, C. P. (1997). Gender presentation: Perceptual and acoustical analyses of voice. *Journal of Voice*, *11*, 307–313.

Ashenfelter, K. T., Boker, S. M., Waddell, J. R., & Vitanov, N. (2009). Spatiotemporal symmetry and multifractal structure of head movements during dyadic conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1072–1091.

Azul, D., & Hancock, A. (2018). Theory-guided, person centered voice and communication support for gender diverse people. In *World Professional Association for Transgender Health (WPATH) Symposium*. Buenos Aires.

Baken, R., & Orlikoff, R. (2000). *Clinical measurement of speech and voice* (2nd ed.). San Diego: Singular Publishing Group.

Bauer, G., & Scheim, A. (2015). *Transgender people in Ontario, Canada: Statistics from the Trans PULSE Project to inform human rights policy*. London, ON.

Beemyn, G., & Rankin, S. (2011). *The lives of transgender people*. New York: Columbia University Press.

Bellaire, K., Yorkston, K. M., & Beukelman, D. R. (1986). Modification of breath patterning to increase naturalness of a mildly dysarthric speaker. *Journal of Communication Disorders*, *19*(4), 271–280. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3745488

Benjamin, H. (1966). *The transsexual phenomenon*. New York: Julian Press.

Bente, G., Donaghy, W. C., & Suwelack, D. (1998). Sex differences in body movement and visual attention: An integrated analysis of movement and gaze in mixed-sex dyads. *Journal of Nonverbal Behavior*, *22*(1), 31–58.

Bhuta, T., Patrick, L., & Garnett, J. D. (2004). Perceptual evaluation of voice quality and its correlation with acoustic measurements. *Journal of Voice*, *18*(3), 299–304. http://doi.org/10.1016/j.jvoice.2003.12.004

Bodoin, E. M., Byrd, C. T., & Adler, R. K. (2014). The clinical profile of the male-to-female transgender person of the 21st century. *Contemporary Issues in Communication Science and Disorders*, *41*, 39–54.

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer [Computer program] Version 6.0.26. Retrieved March 2, 2017, from http://www.praat.org/

Boker, S. M., Cohn, J. F., Theobald, B.-J., Matthews, I., Mangini, M., Spies, J. R., … Brick, T. R. (2011). Something in the way we move: Motion dynamics, not perceived sex, influence head movements in conversation. *Journal of Experimental Psychology. Human Perception and Performance*, *37*, 874–891. http://doi.org/10.1037/a0021928

Boonin, J. (2006a). Articulation. In R. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide* (pp. 225–236). San Diego: Plural Publishing Inc.

Boonin, J. (2006b). Rate and volume. In R. K. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide* (pp. 237–251). San Diego, CA: Plural Publishing Inc.

Boonin, J. (2012). Rate and volume. In R. K. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide* (Second, pp. 237–251). San Diego, CA: Plural Publishing Inc.

Briton, N. J., & Hall, J. A. (1995). Beliefs About female and male nonverbal communication. *Sex Roles*, *32*(1/2), 79–90.

Brockmann, M., Drinnan, M. J., Storck, C., & Carding, P. N. (2011). Reliable jitter and shimmer measurements in voice clinics: The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *Journal of Voice*, *25*(1), 44–53. http://doi.org/10.1016/j.jvoice.2009.07.002

Brown, M. L., & Rounsley, C. A. (2003). *Trueselves: Understanding transsexualism for families, friends, coworkers, and helping professionals*. San Francisco: Jossey-Bass.

Burns-Ardolino, W. A. (2003). Reading woman: Displacing the foundations of femininity. *Hypatia*, *18*(3), 42–59.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. New York: Routledge.

Butler, J. (2004). *Undoing gender*. New York: Routledge.

Byrne, L. A. (2007). *My life as a woman: Placing communication within the social context of life for transsexual women*. La Trobe University.

Byrne, L. A., Dacakis, G., & Douglas, J. M. (2003). Self-perceptions of pragmatic communication abilities in male-to-female transsexuals. *International Journal of Speech-Language Pathology*, *5*(1), 15–25. http://doi.org/10.1080/14417040510001669011

Carew, L., Dacakis, G., & Oates, J. (2007). The effectiveness of oral resonance therapy on the perception of femininity of voice in male-to-female transsexuals. *Journal of Voice*, *21*(5), 591–603. http://doi.org/10.1016/j.jvoice.2006.05.005

Carney, D. R., Cuddy, A. J. C., & Yap, A. J. (2010). Power posing: Brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychological Science*, *21*(10), 1363–1368. http://doi.org/10.1177/0956797610383437

Coleman, E., Bockting, W., Botzer, M., Cohen-Kettenis, P., DeCuypere, G., Feldman, J., … Zucker, K. (2011). Standards of Care for the Health of Transsexual, Transgender, and Gender-Nonconforming People, Version 7. *International Journal of Transgenderism*, *13*(4), 165–232. http://doi.org/10.1080/15532739.2011.700873

Coleman, R. (1983). Acoustic correlates of speaker sex identification: Implications for the transsexual voice. *The Journal of Sex Research*, *19*(3), 293–306. Retrieved from http://www.tandfonline.com/doi/pdf/10.1080/00224498309551189

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*(1), 168–180.

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (n.d.). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement.

Conron, K. J., & Scott, G. (2012). Transgender health in Massachusetts : Results from a household probability sample of adults. *American Journal of Public Health*, *102*(1), 118–122. http://doi.org/10.2105/AJPH.2011.300315

Cruice, M., Worrall, L., & Hickson, L. (2000). Quality-of-life measurement in speech pathology and audiology Quality-of-life in speech pathology and audiology. *Asia Pacific Journal of Speech, Language and Hearing*, *5*(1), 1–20. http://doi.org/10.1179/136132800807547546

Dacakis, G. (2002). The role of voice therapy in male-to-female transsexuals. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *10*, 173–177. http://doi.org/10.1097/01.MOO.0000014709.20032.9B

Dacakis, G., & Davies, S. (2012). Transsexual Voice Questionnaire (male-to-female).

Dacakis, G., Davies, S., Oates, J. M., Douglas, J. M., & Johnston, J. R. (2013). Development and preliminary evaluation of the Transsexual Voice Questionnaire for male-to-female transsexuals. *Journal of Voice : Official Journal of the Voice Foundation*, *27*(3), 312–20. http://doi.org/10.1016/j.jvoice.2012.11.005

Dacakis, G., Oates, J., & Douglas, J. (2012). Beyond voice: Perceptions of gender in male-to-female transsexuals. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *20*(3), 165–170. http://doi.org/10.1097/MOO.0b013e3283530f85

Dacakis, G., Oates, J., & Douglas, J. (2017a). Associations between the Transsexual Voice Questionnaire (TVQ MtF) and self-report of voice femininity and acoustic voice measures. *International Journal of Language & Communication Disorders*, *52*(6), 831–838. http://doi.org/10.1111/1460-6984.12319

Dacakis, G., Oates, J. M., & Douglas, J. M. (2017b). Exploring the validity of the Transsexual Voice Questionnaire (male-to-female): Do TVQ scores differentiate between MtF women who have had gender reassignment surgery and those who have not? *International Journal of Transgenderism*, *17*(3–4), 124–130. http://doi.org/10.1080/15532739.2016.1222922

Dacakis, G., Oates, J. M., & Douglas, J. M. (2017c). Further evidence of the construct validity of the Transsexual Voice Questionnaire (TVQ MtF) using principal components analysis. *Journal of Voice*, *31*(2), 142–148. http://doi.org/10.1016/j.jvoice.2016.07.001

Dacakis, G., Pasricha, N., & Oates, J. (2005). Functional Communication Satisfaction Questionnaire for Male-to-Female Transsexuals. La Trobe University.

Dahl, K. (2018). *Correlating Speech and Voice Features of Transgender Women with Ratings of Femininity and Gender*. University of Rhode Island. Retrieved from http://digitalcommons.uri.edu/cgi/viewcontent.cgi?article=2229&context=theses

Davies, S. M., & Johnston, J. R. (2015). Exploring the validity of the Transsexual Voice Questionnaire. *Canadian Journal of Speech-Language Pathology and Audiology*, *39*(1), 40–51.

Davies, S., Papp, V. G., & Antoni, C. (2015). Voice and communication change for gender nonconforming individuals: Giving voice to the person inside. *International Journal of Transgenderism*, *16*(3), 117. http://doi.org/10.1080/15532739.2015.1075931

Davis, A. (1969). *In the pink of the night*. Metro-Goldwyn-Mayer.

De Cuypere, G., Van Hemelrijck, M., Michel, a, Carael, B., Heylens, G., Rubens, R., … Monstrey, S. (2007). Prevalence and demography of transsexualism in Belgium. *European Psychiatry : The Journal of the Association of European Psychiatrists*, *22*(3), 137–41. http://doi.org/10.1016/j.eurpsy.2006.10.002

Eadie, T. L., & Doyle, P. C. (2002). Direct magnitude estimation and interval scaling of naturalness and severity in tracheoesophageal (TE) speakers. *Journal of Speech, Language, and Hearing Research*, *45*(6), 1088–1096. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12546479

Eadie, T. L., Lamvik, K., Baylor, C. R., Yorkston, K. M., Kim, J., & Amtmann, D. (2014). Communicative participation and quality of life in head and neck cancer. *Annals of Otology, Rhinology & Laryngology*, *123*(4), 257–264. http://doi.org/10.1177/0003489414525020

Ferrans, C. E., & Powers, M. J. (1985). Quality of Life Index: Development and psychometric properties. *Advances in Nursing Science*, *8*(1), 15–24.

Ferrans, C. E., & Powers, M. J. (1998). Quality of life index generic version - III. Retrieved from https://www.uic.edu/orgs/qli/questionaires/pdf/genericversionIII/generic.pdf

Fitzsimons, M., Sheahan, N., & Staunton, H. (2001). Gender and the integration of acoustic dimensions of prosody: implications for clinical studies. *Brain and Language*, *78*(1), 94–108. http://doi.org/10.1006/brln.2000.2448

Flores, A. R., Herman, J. L., Gates, G. J., & Brown, T. N. T. (2016). *How many adults identify as transgender in the United States?* Los Angeles, CA.

Fox, C. M., & Boliek, C. A. (2012). Intensive voice treatment (LSVT LOUD) for children with spastic cerebral palsy and dysarthria. *Journal of Speech, Language, and Hearing Research*, *55*(3), 930–45. http://doi.org/10.1044/1092-4388(2011/10-0235)

Gagné, P., & Tewksbury, R. (1999). Knowledge and power, body and self: An analysis of knowledge systems and the transgendered self. *The Sociological Quarterly*, *40*(1), 59–83. http://doi.org/10.1111/j.1533-8525.1999.tb02358.x

Gelfer, M. P., & Bennett, Q. E. (2013). Speaking fundamental frequency and vowel formant frequencies: Effects on perception of gender. *Journal of Voice*, *27*(5), 556–566. http://doi.org/10.1016/j.jvoice.2012.11.008

Gelfer, M. P., & Mikos, V. a. (2005). The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. *Journal of Voice*, *19*(4), 544–554. http://doi.org/10.1016/j.jvoice.2004.10.006

Gelfer, M. P., & Mordaunt, M. (2012). Pitch and intonation. In R. K. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and communication therapy for the transgender/transsexual Client: A comprehensive clinical guide* (Second Edi, pp. 187–224). San Diego, CA: Plural Publishing Inc.

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 Voice*, *14*(1), 22–33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10764114

Gelfer, M. P., & Tice, R. M. (2013). Perceptual and acoustic outcomes of voice therapy for male-to-female transgender individuals immediately after therapy and 15 months later. *Journal of Voice*, *27*(3), 335–347. http://doi.org/10.1016/j.jvoice.2012.07.009

Gelfer, M. P., & Van Dong, B. R. (2013). A preliminary study on the use of vocal function exercises to improve voice in male-to-female transgender clients. *Journal of Voice*, *27*(3), 321–334. http://doi.org/10.1016/j.jvoice.2012.07.008

Gelfer, M. P., & Young, S. R. (1997). Comparisons of intensity measures and their stability in male and female speakers. *Journal of Voice*, *11*(2), 178–186.

Gender Finder. (2017). Edmonton: Webzao Innovations Inc.

Gherovici, P. (2010). *Please select your gender: From the invention of hysteria to the democratizing of transgenderism*. New York: Routledge.

Gillespie, A. I., & Gartner-schmidt, J. (2018). Voice-specialized speech-language pathologist's criteria for discharge from voice therapy. *Journal of Voice*, *32*(3), 332–339. http://doi.org/10.1016/j.jvoice.2017.05.022

Gorham-Rowan, M., & Morris, R. (2006). Aerodynamic analysis of male-to-female transgender voice. *Journal of Voice*, *20*(2), 251–262. http://doi.org/10.1016/j.jvoice.2005.03.004

Grant, J. M., Mottet, L. A., Tanis, J., Min, D., Harrison, J., Herman, J. L., & Keisling, M. (2011). *Injustice at every turn: A report of the national transgender discrimination survey*.

Günzburger, D. (1989). Voice adaptation by transsexuals. *Clinical Linguistics & Phonetics*, *3*(2), 163–172.

Günzburger, D. (1993). An acoustic analysis and some perceptual data concerning voice change in male-female trans-sexuals. *European Journal of Disorders of Communication*, *28*, 13–21.

Günzburger, D. (1995). Acoustic and perceptual implications of the transsexual voice. *Archives of Sexual Behavior*, *24*(3), 339–348. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7611850

Hall, J. A. (1984). *Nonverbal Sex Differences: Communication Accuracy and Expressive Style*. Baltimore: Johns Hopkins University Press.

Hall, J. A. (2006). Nonverbal behavior, status, and gender: How do we understand their relations? *Psychology of Women Quarterly*, *30*(2006), 384–391.

Hall, J. A., & Friedman, G. B. (1999). Status, gender, and nonverbal behavior: A study of structured interactions between employees of a company. *Personality and Social Psychology Bulletin*, *25*(9), 1082–1091. http://doi.org/10.1177/01461672992512002

Hamilton, D. L., Sherman, S. J., & Ruvolo, C. M. (1990). Stereotype-based expectancies: Effects on information processing and social behavior. *Journal of Social Issues*, *46*(2), 35–60.

Hancock, A. B. (2017). An ICF perspective on voice-related quality of life of American transgender women. *Journal of Voice*, *31*(1), 115.e1-115.e8. http://doi.org/10.1016/j.jvoice.2016.03.013

Hancock, A. B., & Garabedian, L. M. (2013). Transgender voice and communication treatment: a retrospective chart review of 25 cases. *International Journal of Language & Communication Disorders*, *48*(1), 54–65. http://doi.org/10.1111/j.1460-6984.2012.00185.x

Hancock, A. B., Krissinger, J., & Owen, K. (2011). Voice perceptions and quality of life of transgender people. *Journal of Voice*, *25*(5), 553–558. http://doi.org/10.1016/j.jvoice.2010.07.013

Hancock, A. B., & Pool, S. F. (2017). Influence of listener characteristics on perceptions of sex and gender. *Journal of Language and Social Psychology*, *36*(5), 599–610. http://doi.org/10.1177/0261927X17704460

Hancock, A., Colton, L., & Douglas, F. (2014). Intonation and gender perception: Applications for transgender speakers. *Journal of Voice*, *28*(2), 203–209. http://doi.org/10.1016/j.jvoice.2013.08.009

Hardy, T. L. D., Boliek, C. A., Aalto, D., Lewicke, J., Wells, K., & Rieger, J. M. (2019). *Contributions of voice and nonverbal communication to perceived masculinity-femininity for cisgender and transgender communicators*. Manuscript submitted for publication.

Hardy, T. L. D., Boliek, C. A., Wells, K., Aalto, D., Lewicke, J. & Rieger, J. M. (2019). *Let me see your body talk: An exploration of gender differences in upper body nonverbal communication behaviors*. Manuscript in preparation.

Hardy, T. L. D., Boliek, C. A., Wells, K., Dearden, C., Zalmanowitz, C., & Rieger, J. M. (2016). Pretreatment acoustic predictors of gender, femininity, and naturalness ratings in individuals with male-to-female gender identity. *American Journal of Speech-Language Pathology*, *25*, 125–137. http://doi.org/10.1044/2015_AJSLP-14-0098

Hardy, T. L. D., Boliek, C. A., Wells, K., & Rieger, J. M. (2013). The ICF and Male-to-Female Transsexual Communication. *International Journal of Transgenderism*, *14*, 196–208. http://doi.org/10.1080/15532739.2014.890561

Hardy, T. L. D., Rieger, J. M., Wells, K., & Boliek, C. A. (2018). Acoustic predictors of gender attribution, masculinity−femininity, and vocal naturalness ratings amongst transgender and cisgender speakers. *Journal of Voice*, 1–16. http://doi.org/10.1016/j.jvoice.2018.10.002

Hardy, T. L. D., Rieger, J. M., Wells, K., & Boliek, C. A. (2019). *Associations between voice and gestural characteristics of transgender women and self-rated femininity, satisfaction, and quality of life.* Manuscript submitted for publication.

Helweg-Larsen, M., Cunningham, S. J., Carrico, A., & Pergram, A. M. (2004). To nod or not to nod: An observational study of nonverbal communication and status in female and male college students. *Psychology of Women Quarterly*, *28*, 358–361.

Henley, N. M. (1977). *Body politics*. Englewood Cliffs, NJ: Prentice Hall.

Henley, N. M. (1995). Body politics revisited: What do we know today? In P. J. Kalbfleish & M. J. Cody (Eds.), *Gender, power, and communication in human relationships* (pp. 27–61). Hillsdale, NJ: Lawrence Erlbaum.

Henton, C. G. (1989). Fact and fiction in the description of female and male pitch. *Language and Communication*, *9*(4), 299–311.

Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology*, *11*, 880–885. http://doi.org/10.1016/S0960-9822(01)00243-3

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *97*(5 Pt 1), 3099–3111. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7759650

Hillenbrand, J. M., & Clark, M. J. (2009). The role of f 0 and formant frequencies in distinguishing the voices of men and women. *Attention, Perception, & Psychophysics*, *71*(5), 1150–1166. http://doi.org/10.3758/APP.71.5.1150

Hirsch, S. (2006). Nonverbal communication: A multicultural view. In R. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide* (pp. 283–317). San Diego: Plural Publishing Inc.

Hirsch, S. (2017). Combining voice, speech science and art approaches to resonant challenges in transgender voice and communication training. *Perspectives of the ASHA Special Interest Groups: SIG 10*, *2*(Part 2), 74–83.

Hirsch, S., & Boonin, J. (2012). Nonverbal communication: Assessment and training. In R. K. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide* (Second, pp. 353–391). San Diego, CA: Plural Publishing Inc.

Hirsch, S., & Van Borsel, J. (2012). Nonverbal communication: A multicultural view. In R. K. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and Communication Therapy for the Transgender/Transsexual Client: A Comprehensive Clinical Guide* (Second, pp. 319–352). San Diego, CA: Plural Publishing Inc.

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

Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life. *Journal of Voice*, *13*(4), 557–569.

Holmberg, E. B., 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. http://doi.org/10.1016/j.jvoice.2009.02.002

Hosmer, D. J., Lemeshow, S., & Sturdivant, R. (2013). Model-building strategies and methods for logistic regression. In *Applied logistic regression* (3rd ed., pp. 89–151). New Jersey: John Wiley & Sons, Inc.

Hostetter, A. B., & Hopkins, W. D. (2002). The effect of thought structure on the production of lexical movements. *Brain and Language*, *82*(1), 22–29. http://doi.org/10.1016/S0093-934X(02)00009-3

Hunt, K. W. (1965). A synopsis of clause-to-sentence length factors. *The English Journal*, *54*(4), 305–309.

International classification of functioning disability and health (ICF). (2001). Geneva, Switzerland: World Health Organization (WHO).

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jaconsen, G., & Benninger, M. S. (1997). The Voice Handicap Index (VHI): Development and validation. *American Journal of Speech-Language Pathology*, *6*(3), 66–70.

James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016a). *Executive summary of the report of the 2015 U.S. transgender survey*. Washington, DC.

James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016b). *The report of the 2015 U.S. transgender survey*. Washington, DC.

Johnson, K. L., McKay, L. S., & Pollick, F. E. (2011). He throws like a girl (but only when he's sad): Emotion affects sex-decoding of biological motion displays. *Cognition*, *119*(2), 265–80. http://doi.org/10.1016/j.cognition.2011.01.016

Jones, E., Gallois, C., Callan, V., & Barker, M. (1999). Strategies of accommodation:: Development of a coding system for conversational interaction. *Journal of Language and Social Psychology*, *18*(June), 123–151. http://doi.org/10.1177/0261927X99018002001

KayPentax. (2008). Multi-Speech Software. Lincoln Park, NJ: KayPENTAX.

Kazdin, A. E., Sherick, R. B., Esveldt-Dawson, K., & Rancurello, M. D. (1985). Nonverbal behavior and childhood depression. *Journal of the American Academy of Child Psychiatry*, *24*(3), 303–309. http://doi.org/10.1016/S0002-7138(09)61091-8

Kelley, M. (2016). Boliek Lab Segmental Variables Extractor. Edmonton.

Kelley, M. (2017). Fo Script. Edmonton.

Kennedy, C. W., & Camden, C. (1984). Interruptions and nonverbal gender differences. *Journal of Nonverbal Behavior*, *8*(2), 91–108. http://doi.org/10.1007/BF00986997

King, R. S., Brown, G. R., & McCrea, C. R. (2012). Voice parameters that result in identification or misidentification of biological gender in male-to-female transgender veterans. *International Journal of Transgenderism*, *13*(3), 117–130. http://doi.org/10.1080/15532739.2011.664464

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, *87*(2), 820–857. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2137837

Koffi, E. (2018). A just noticeable difference (JND) reanalysis of Fry's original acoustic correlates of stress in American English. *Linguistic Portfolios*, *7*(2), 1–25.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. http://doi.org/10.1016/j.jcm.2016.02.012

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, *21*(6), 575–580.

Krause, W. (2006). Androgens in the demography of male life course--A review. *Social Biology*, *53*(1–2), 4–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21516946

Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research*, *36*(1), 21–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8450660

LaFrance, M., Hecht, M. a., & Paluck, E. L. (2003). The contingent smile: A meta-analysis of sex differences in smiling. *Psychological Bulletin*, *129*(2), 305–334. http://doi.org/10.1037/0033-2909.129.2.305

Laqueur, T. (1990). *Making sex: Body and gender from the Greeks to Freud*. Cambridge, MA: Harvard University Press.

Leung, Y., Oates, J., & Pang Chan, S. (2018). Voice, articulation, and prosody contribute to listener perceptions of speaker gender: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, *61*, 266–297. http://doi.org/10.1044/2017_JSLHR-S-17-0067

Martins, R. H. G., Tavares, E. L. M., & Pessin, A. B. B. (2017). Are vocal alterations caused by smoking in Reinke's Edema in women entirely reversible after microsurgery and smoking cessation? *Journal of Voice*, *31*(3), 380.e11-380.e14. http://doi.org/10.1016/j.jvoice.2016.06.012

Mather, G., & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society B: Biological Sciences*, *258*, 273–279. Retrieved from http://rspb.royalsocietypublishing.org/

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

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. http://doi.org/10.1016/j.jvoice.2006.12.010

Mendoza, E., Valencia, N., Muñoz, J., & Trujillo, H. (1996). Differences in voice quality between men and women: Use of the long-term average spectrum (LTAS). *Journal of Voice*, *10*(1), 59–66. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8653179

Michael-Michelke, D. (1997). *Ratings of speech naturalness in persons with dysphonia*. University of Minnesota.

Milenkovic, P. H. (2001). TF32. Madison, WI: University of Wisconsin-Madison.

Mordaunt, M. (2006). Pitch and intonation. In R. K. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and communication therapy for the transgender/transsexual Client: A comprehensive clinical guide* (pp. 168–208). San Diego: Plural Publishing Inc.

Mount, K. H., & Salmon, S. J. (1988). Changing the vocal characteristics of a postoperative transsexual patient: A longitudinal study. *Journal of Communication Disorders*, *21*(3), 229–238. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3417881

Mundt, J. C., Snyder, P. J., Cannizzaro, M. S., Chappie, K., & Geralts, D. S. (2007). Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. *Journal of Neurolinguistics*, *20*, 50–64. http://doi.org/10.1016/j.jneuroling.2006.04.001

Murad, M. H., Elamin, M. B., Garcia, M. Z., Mullan, R. J., Erwin, P. J., & Montori, V. M. (2010). Hormonal therapy and sex reassignment : A systematic review and meta-analysis of quality of life and psychosocial outcomes. *Clinical Endocrinology*, *72*, 214–231. http://doi.org/10.1111/j.1365-2265.2009.03625.x

Nagoshi, J. L., & Brzuzy, S. /I. (2010). Transgender theory: Embodying research and practice. *Affilia*, *25*(4), 431–443. http://doi.org/10.1177/0886109910384068

Nicoladis, E., Pika, S., & Marentette, P. (2009). Do French – English Bilingual children gesture more than monolingual children? *Journal of Psycholinguistic Research*, *38*, 573–585. http://doi.org/10.1007/s10936-009-9121-7

Nicoladis, E., Pika, S., Yin, H., & Marentette, P. (2007). Gesture use in story recall by Chinese–English bilinguals. *Applied Psycholinguistics*, *28*, 721–735. http://doi.org/10.1017/S0142716407070385

Nilsonne, Å., Sundberg, J., Ternström, S., & Askenfelt, A. (1988). Measuring the rate of change of voice fundamental frequency in fluent speech during mental depression. *The Journal of the Acoustical Society of America*, *83*, 716–728. http://doi.org/10.1121/1.396114

Nisltio, M., Tanaka, Y., & Sakabibara, C. (2007). Effectiveness of speech rate-conversion software for patients with dysarthria. *Journal of Communications Research*, *2*(1), 1–12.

Norton, A. (2000). Voice therapy helps transsexuals. Retrieved February 19, 2015, from https://groups.yahoo.com/neo/groups/TNUKdigest/conversations/messages/869

Nusbaum, H. C., Francis, A. L., & Henly, A. S. (1995). Measuring the naturalness of synthetic speech. *International Journal of Speech Technology*, *1*, 7–19.

Oates, J., & Dacakis, G. (2015). Transgender voice and communication: Research evidence underpinning voice intervention for male-to-female transsexual women. *Perspectives on Voice and Voice Disorders*, *25*(2), 48–58. http://doi.org/10.1044/vvd25.2.48

Oates, J. M., & Dacakis, G. (1983). Speech pathology considerations in the management of transsexualism—A review. *British Journal of Disorders of Communication*, *18*(3), 139–151. http://doi.org/10.3109/13682828309012237

Oates, J. M., & Dacakis, G. (1997). Voice change in transsexuals. *Venereology*, *10*(3), 178–187.

Owen, K., & Hancock, A. B. (2010). The role of self- and listener perceptions of femininity in voice therapy. *International Journal Of Transgenderism*, *12*, 272–284. http://doi.org/10.1080/15532739.2010.550767

Parrish, W. M. (1951). The concept of "naturalness." *Quarterly Journal of Speech*, *37*, 448–454.

Pasricha, N., Dacakis, G., & Oates, J. (2008). Communicative satisfaction of male-to-female transsexuals. *Logopedics Phoniatrics Vocology*, *33*, 25–34. http://doi.org/10.1080/14015430701514500

Peterson, P. (1976). An investigation of sex differences in regard to nonverbal body gestures. In B. Eakins, G. Eakins, & B. Lieb-Brilhart (Eds.), *SISCOM '75: Women's (and Men's) Communication* (pp. 20–27). Falls Church, VA.

Pollick, F. E., Kay, J. W., Heim, K., & Stringer, R. (2005). Gender recognition from point-light walkers. *Journal of Experimental Psychology. Human Perception and Performance*, *31*(6), 1247–65. http://doi.org/10.1037/0096-1523.31.6.1247

Pollick, F. E., Lestou, V., Ryu, J., & Cho, S.-B. (2002). Estimating the efficiency of recognizing gender and affect from biological motion. *Vision Research*, *42*, 2345–2355. http://doi.org/10.1016/S0042-6989(02)00196-7

Porter, C. C. (2012). *Voice quality and gender identification: Acoustic and perceptual analysis*. Dalhousie University.

Prosser, J. (1998). *Second skins: The body narratives of transsexuality*. New York: Columbia University Press.

Ratcliff, A., Coughlin, S., & Lehman, M. (2002). Factors influencing ratings of speech naturalness in augmentative and alternative communication. *AAC Augmentative and Alternative Communication*, *18*, 11–19.

Raymond, J. (1979). *The transsexual empire: The making of the she-male*. Boston: Beacon Press.

Rehman, J., Lazer, S., Benet, A. E., Schaefer, L., & Melman, A. (1999). The reported sex and surgery satisfactions of 28 postoperative male-to-female transsexual patients. *Archives of Sexual Behavior*, *28*(1), 71–89.

Richards, C., Bouman, W. P., Seal, L., Barker, M. J., Nieder, T. O., Sjoen, G. T., … Barker, M. J. (2016). Non-binary or genderqueer genders. *International Review of Psychiatry*, *28*(1), 95–102. http://doi.org/10.3109/09540261.2015.1106446

Robb, M. P., Maclagan, M. A., & Chen, Y. (2004). Speaking rates of American and New Zealand varieties of English Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics & Phonetics*, *18*(1), 1–15. http://doi.org/10.1080/0269920031000105336

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, *112*(4), 585–615.

Samons, S. L. (2009). *When the opposite sex isn't: Sexual orientation in male-to-female transgender people*. New York: Alyson Books.

Sánchez, F. J., Greenberg, S. T., Liu, W. M., & Vilain, E. (2009). Reported effects of masculine ideals on gay men. *Psychology of Men & Masculinity*, *10*(1), 73–87. http://doi.org/10.1037/a0013513.Reported

Shelley, C. A. (2008). *Transpeople: Repudiation, trauma, healing*. Toronto: University of Toronto Press.

Skuk, V. G., & Schweinberger, S. R. (2014). Influences of fundamental frequency, formant frequencies, aperiodicity, and spectrum level on the perception of voice gender. *Journal of Speech, Language, and Hearing Research*, *57*, 285–296. http://doi.org/10.1044/1092-4388(2013/12-0314)(VTL)

Smithson, L., & Nicoladis, E. (2014). Lending a hand to imagery? The impact of visuospatial working memory interference upon iconic gesture production in a narrative task. *Journal of Nonverbal Behavior*, *38*(2), 247–258. http://doi.org/10.1007/s10919-014-0176-2

Snow, M. P., & Williges, R. C. (1998). Empirical models based on free-modulus magnitude estimation of perceived presence in virtual environments. *Human Factors and Ergonomics Society*, *40*(3), 1–14.

Södersten, M., Hertegård, S., & Hammarberg, B. (1995). Glottal closure, transglottal airflow, and voice quality in healthy middle-aged women. *Journal of Voice*, *9*(2), 182–197. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7620541

Spade, D. (2006). Mutilating gender. In S. Stryker & S. Whittle (Eds.), *The transgender studies reader* (pp. 315–332). New York: Routledge.

Spencer, L. E. (1988). Speech characteristics of male-to-female transsexuals: A perceptual and acoustic study. *Folia Phoniatrica*, *40*(1), 31–42.

Stevens, S. S. (1975). *Psychophysics*. (G. Stevens, Ed.). New York, NY: John Wiley & Sons, Inc.

Stone, S. (2006). The empire strikes back: A posttranssexual manifesto. In S. Stryker & S. Whittle (Eds.), *The transgender studies reader* (pp. 221–235). New York: Routledge.

Stryker, S. (2008). *Transgender history*. Berkeley, CA: Seal Press.

T'Sjoen, G., Moerman, M., Van Borsel, J., Feyen, E., Rubens, R., Monstrey, S., … De Cuypere, G. (2006). Impact of voice in transsexuals. *International Journal Of Transgenderism*, *9*(1), 1–7. http://doi.org/10.1300/J485v09n01

Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics*. Boston: Pearson.

Terango, L. (1966). Pitch and duration characteristics of the oral reading of males on a masculinity-femininity dimension. *Journal of Speech and Hearing Research*, *9*, 590–595.

Teshima, S., Langevin, M., Hagler, P., & Kully, D. (2010). Post-treatment speech naturalness of comprehensive stuttering program clients and differences in ratings among listener groups. *Journal of Fluency Disorders*, *35*(1), 44–58. http://doi.org/10.1016/j.jfludis.2010.01.001

Troje, N. F. (2002a). Gender and personality in human gait. *TEAP Tagung Experimentell Arbeitender Psychologen Chemnitz Germany*.

Troje, N. F. (2002b). The little difference: Fourier based gender classification from biological motion. *Dynamic Perception R P Wrtz and M Lappe Eds Aka Press Berlin*, 115–120.

Tsao, Y. C., & Weismer, G. (1997). Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component. *Journal of Speech and Hearing Research*, *40*(4), 858–866.

Tsao, Y., Weismer, G., & Iqbal, K. (2006). Interspeaker variation in habitual speaking rate : Additional evidence, *49*, 1156–1165.

Tskhay, K. O., & Rule, N. O. (2013). Accuracy in categorizing perceptually ambiguous groups: A review and meta-analysis. *Personality and Social Psychology Review*, *17*(1), 72–86. http://doi.org/10.1177/1088868312461308

Valocchi, S. (2005). Not yet queer enough: The lessons of queer theory for the sociology of gender and sexuality. *Gender & Society*, *19*(6), 750–770. http://doi.org/10.1177/0891243205280294

Van Borsel, J., De Cuypere, G., & Van den Berghe, H. (2001). Physical appearance and voice in male-to-female transsexuals. *Journal of Voice*, *15*(4), 570–575. http://doi.org/10.1016/S0892-1997(01)00059-5

Van Borsel, J., & De Maesschalck, D. (2008). Speech rate in males, females, and male-to-female transsexuals. *Clinical Linguistics & Phonetics*, *22*(9), 679–685. http://doi.org/10.1080/02699200801976695

Van Borsel, J., & Eeckhout, H. (2008). The speech naturalness of people who stutter speaking under delayed auditory feedback as perceived by different groups of listeners. *Journal of Fluency Disorders*, *33*(3), 241–51. http://doi.org/10.1016/j.jfludis.2008.06.004

Van Borsel, J., Janssens, J., & De Bodt, M. (2009). Breathiness as a feminine voice characteristic: A perceptual approach. *Journal of Voice*, *23*(3), 291–294. http://doi.org/10.1016/j.jvoice.2007.08.002

van der Zwan, R., Machatch, C., Kozlowski, D., Troje, N. F., Blanke, O., & Brooks, A. (2009). Gender bending: Auditory cues affect visual judgements of gender in biological motion displays. *Experimental Brain Research*, *198*, 373–382.

Verhoeven, J., De Pauw, G., & Kloots, H. (2004). Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech*, *47*(3), 297–308. http://doi.org/10.1177/00238309040470030401

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., & Vandenbroucke, J. P. (2007). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PLoS Med*, *4*(10), e296. http://doi.org/10.1371/journal.pmed. 0040296 Received:

Vrugt, A., & Kerkstra, A. (1984). Sex differences in nonverbal communication. *Semiotica*, *50*(1/2), 1–41.

Wendahl, R. W. (1966a). Laryngeal analog synthesis of jitter and shimmer auditory parameters of harshness. *Folia Phoniatrica*, *18*, 98–108.

Wendahl, R. W. (1966b). Some parameters of auditory roughness. *Folia Phoniatrica et Logopaedica*, *18*, 26–32.

Whitehill, T. L., Lee, A. S. Y., & Chun, J. C. (2002). Direct magnitude estimation and interval scaling of hypernasality. *Journal of Speech, Language, and Hearing Research*, *45*(1), 80–88. http://doi.org/10.1044/1092-4388(2002/006)

Whittle, S. (2006). Where did we go wrong? Feminism and trans theory - Two teams on the same side? In S. Stryker & S. Whittle (Eds.), *The transgender studies reader* (pp. 194–202). New York: Routledge.

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

Wöllner, C., & Deconinck, F. J. A. (2013). Gender recognition depends on type of movement and motor skill. Analyzing and perceiving biological motion in musical and nonmusical tasks. *Acta Psychologica*, *143*(1), 79–87. http://doi.org/10.1016/j.actpsy.2013.02.012

Yang, P. (2010). Nonverbal gender differences : examining gestures of university-educated Mandarin Chinese speakers *. *Text & Talk*, *30*(3), 333–357.

Zeig, S. (1985). The actor as activator: Deconstructing gender through gesture. *Feminist Issues*, *Spring*, 21–25.

APPENDIX A – ETHICS APPROVAL

Approval Form

Date:	February 5, 2016
Study ID:	Pro00060133
Principal Investigator:	Carol Boliek
Study Title:	Talking the talk and walking the walk: Communication and quality of life in the male-to-female transgender population
Approval Expiry Date:	Friday, February 3, 2017

Approved Consent Form:	Approval Date	Approved Document
	2/5/2016	Cisgender Communicator Consent
	2/5/2016	Rater Consent
	2/5/2016	Transgender Communicator Consent

Sponsor/Funding Agency:	Alberta Innovates Health Solutions	AIHS	Canada
Sponsor/Funding Agency:	SSHRC Doctoral Award		

RSO-Managed Funding:

	Project ID	Project Title	Speed Code	Other Information
View	RES0026946	SAL Talking the talk and walking the walk: Communication and quality of life in the male-to-female transgender population		

Thank you for submitting the above study to the Health Research Ethics Board - Health Panel . Your application, including the following, has been reviewed and approved on behalf of the committee;

- Transgender Communicator Recruitment Poster (1/22/2016)
- Rater Recruitment Poster (2/5/2016)
- Cisgender Communicator Recruitment Poster (1/22/2016)
- Invitation Letter Active Wait Lit (1/26/2016)
- Invitation Letter for Participants who have had Previous Treatment (1/26/2016)
- Protocol - Communicators (1/15/2016)
- Protocol - Raters (1/15/2016)
- Speaker Perception QOL and Comm Sat (1/15/2016)
- Transsexual Voice Questionnaire -= Male-Female (1/26/2016)
- Transgender Subjective Ratings (1/22/2016)
- Communicator Debriefing Script (1/22/2016)
- Rater Debriefing Script (1/22/2016)
- Cisgender Demographics Data Sheet (1/22/2016)
- Transgender Communicator Demographics Data Sheet (1/22/2016)
- Rater Demographics Data Sheet (1/22/2016)
- Quality of Life Index (1/25/2016)

The Health Research Ethics Board assessed all matters required by section 50(1)(a) of the Health Information Act. It has been determined that the recruitment portion of the research described in the ethics application is a retrospective chart review for which subject consent for access to personally identifiable health information would not be reasonable, feasible or practical. Subject consent therefore is not required for access to personally identifiable health information described in the ethics application.

In order to comply with the Health Information Act, a copy of the approval form is being sent to the Office of the Information and Privacy Commissioner.

A renewal report must be submitted next year prior to the expiry of this approval if your study still requires ethics approval. If you do not renew on or before the renewal expiry date, you will have to re-submit an ethics application.

Approval by the Health Research Ethics Board does not encompass authorization to access the patients, staff or resources of Alberta Health Services or other local health care institutions for the purposes of the research. Enquiries regarding Alberta Health Services approvals should be directed to (780) 407-6041. Enquiries regarding Covenant Health should be directed to (780) 735-2274.

Sincerely,

Anthony S. Joyce, Ph.D.
Chair, Health Research Ethics Board - Health Panel

Note: This correspondence includes an electronic signature (validation and approval via an online system).

Transgender Women Communicators

Transfeminine Volunteers Needed

Communication Factors Related to Being Read as a Woman

Study Title:

Talking the talk and walking the walk: Communication and quality of life in the male-to-female transgender population

Purpose:

We are trying to find out which communication patterns are most important in making a person seem to be a man or woman (or neither) to other people and how those factors are related to quality of life

We are looking for:

- Individuals who identify as transgender, transsexual, or trans*
- Individuals who have been living as female for the majority of the time for at least 6 months
- Fluent English speakers
- Individuals who have full movement of their upper body
- Individuals who are free of neurogenic disorders (e.g., Parkinson's Disease, aphasia, dysarthria)

How much time?

- 1 lab visit
- Visit takes 1 - 1½ hours

Where do I go?

Syncrude Centre for Motion and Balance at the Glenrose Rehabilitation Hospital

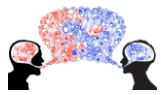
What will I do?

Retell the story of a cartoon while being recorded; fill out questionnaires

Teresa Hardy (Boliek & Rieger Labs): (780) 492-7256
teresa.hardy@ualberta.ca

Cisgender Communicators

Research Volunteers Needed

Communication Characteristics of Story Retell 

Purpose:
We are trying to find out how different people retell stories and how they sound to others

We are looking for:

- Fluent English speakers
- Individuals who have full movement of their upper body
- Individuals who are free of neurogenic disorders (e.g., Parkinson's Disease, aphasia, dysarthria)

How much time?

- 1 lab visit
- Visit takes 1 - 1½ hours

Raters

Where do I go?

Syncrude Centre for Motion and Balance at the Glenrose Rehabilitation Hospital

What will I do?

Perform three short speaking tasks (e.g., retell the story of a short cartoon) while being recorded

Teresa Hardy (Boliek & Rieger Labs): (780) 492-7256
teresa.hardy@ualberta.ca

Raters

Research Volunteers Needed

Gender Differences in Story Retell 

Purpose:
We are trying to find out which communication patterns are most important in making a person seem to be a man or woman (or neither) to other people

We are looking for:

- Fluent English speakers
- Individuals who are free from hearing or uncorrected visual difficulties
- Individuals who have no special training in listening to voices or rating gestures
- Individuals who do not have language, learning, or cognitive disabilities

How much time?

- 1 lab visit
- Visit takes about 2.5 hours

Where do I go?

Clinical Sciences Building at the University of Alberta

What will I do?

Make ratings about the voices and gestures of different people describing a cartoon

Call Teresa (Boliek Lab): (780) 492-7256
cblab@ualberta.ca

Ethics PRO 00060133

APPENDIX C – RECRUITMENT LETTERS

Active Wait List

PATIENT ADDRESS

Date, 2016

Dear _____,

I hope that this letter finds you well and in good health.

I am writing to you today in the hopes that you will be willing to participate in a new study, titled:

Talking the talk and walking the walk: Communication and quality of life in the male-to-female transgender population.

The major aim of this study is to identify the most important parts of communication in making a person come across as a woman or a man. This information will help clinicians create better treatment programs for people who are pursuing gender transition and want to make their communication more feminine. This study is part of Teresa Hardy's PhD work. Teresa is a speech-language pathologist and a clinician in the voice program at the Glenrose Rehabilitation Hospital. The lead on this study is Dr. Carol Boliek. She is a professor at the University of Alberta.

I identified your name from the treatment database for the voice program at the Glenrose Rehabilitation Hospital. I typically have access to this information in my role as a clerk for the Communication Disorders service. I am not part of the research team and your name has not been and will not be released to them. I am forwarding this letter on their behalf.

If you would like to learn more about the study and/or would like to know if you can participate in some way, please contact **Dr. Carol Boliek** at **(780) 492-0841** or **carol.boliek@ualberta.ca.**

Thank you for your time in reading this letter.

Sincerely,

Donna Sheptycki
On behalf of the Communication Feminization Research Team

Previously Received Training

PATIENT ADDRESS

Date, 2016

Dear _____,

I hope that this letter finds you well and in good health.

I am writing to you today in the hopes that you will be willing to participate in a new study, titled:

Talking the talk and walking the walk: Communication and quality of life in the male-to-female transgender population.

The major aim of this study is to identify the most important parts of communication in making a person come across as a woman or a man. This information will help clinicians create better treatment programs for people who are pursuing gender transition and want to make their communication more feminine. This study is part of Teresa Hardy's PhD work. Teresa is a speech-language pathologist and a clinician in the voice program at the Glenrose Rehabilitation Hospital. The lead on this study is Dr. Carol Boliek. She is a professor at the University of Alberta.

I identified your name from the treatment database for the voice program at the Glenrose Rehabilitation Hospital. I typically have access to this information in my role as a clerk for the Communication Disorders service. I am not part of the research team and your name has not been and will not be released to them. I am forwarding this letter on their behalf.

If you would like to learn more about the study and/or would like to know if you can participate in some way, please contact **Teresa Hardy** at **(780) 492-7256** or **teresa.hardy@ualberta.ca.**

Thank you for your time in reading this letter.

Sincerely,

Donna Sheptycki
On behalf of the Communication Feminization Research Team

APPENDIX D – PARTICIPANT INFORMATION LETTERS/CONSENT FORMS

Cisgender Communicators

PARTICIPANT CONSENT FORM

Title of Study: Communication Characteristics of Story Retell

Principal Investigator: Carol Boliek, PhD

Research/Study Coordinator: Teresa Hardy, Speech-Language Pathologist, PhD Candidate

---

Why am I being asked to take part in this research study?
You are being asked to be in this study because you do not have any disorders that may impact your ability to communicate. The way a person communicates is a very important part of their identity and impacts the way others perceive them. We want to understand better the communication-based characteristics of story retelling and how others perceive those characteristics. This information will help us to provide better treatment for people who are having difficulty with their communication.

Before you make a decision one of the researchers will go over this form with you. You are encouraged to ask questions if you feel anything needs to be made clearer. You will be given a copy of this form for your records.

What is the reason for doing the study?
We want to find out what aspects of a person's communication are most meaningful to the people with whom they interact and how these characteristics influence the way a person comes across to others.

This study is being completed as part of a graduate thesis.

What will I be asked to do?
You will be asked to watch the Pink Panther cartoon, "In the Pink of the Night" and then recount the story to a researcher while being recorded. The visit should take 1 to 1.5 hours and you will only need to attend once. The study should take about one year to complete once all data are collected.

Information Sharing and Consent
When you arrive for your visit, a researcher will explain the study and what you will be asked to do. You then will be asked to sign a consent form to show that you agree to participate in the study, if that is the case.

Story Retell (i.e., Cartoon Description) Task
Once you arrive at the Syncrude Centre for Motion and Balance at the Glenrose Rehabilitation Hospital, you will be fitted with approximately 12 passive motion markers and a small microphone. The passive motion markers resemble small Styrofoam balls and are attached to the head, sternum, and upper body joints using adhesive pads. The microphone will be fixed to your forehead to keep it at a constant distance from your mouth. You may be provided with a standardized shirt or smock and hair band (if necessary) to allow for proper placement of the equipment. Next, you will be seated on a raised stool and you will watch the Pink Panther cartoon, "In the Pink of the Night". You will be asked to retell the events of the cartoon in as much detail as possible. Your voice will be recorded through the microphone and your movements will be recorded by a group of video cameras placed around the room, much like is done when creating computer generated imagery (CGI) movies or video games. As a result, the video recording will not show your appearance; rather, it will look like a moving stick figure. You also will be recorded repeating or reading three short phrases and holding out a vowel sound. When you have finished these communication tasks, your participation in the study will be complete.

What are the risks and discomforts?
There is a small chance that you may become fatigued while recalling the cartoon. You also may feel mild awkwardness when being fitted with the recording equipment. There are no other known risks associated with this study; however, it is not possible to know all of the risks that may happen in a study. The researchers have taken all reasonable safeguards to minimize any known risks to a study participant.

What are the benefits to me?
You are not expected to get any benefit from being in this research study; however, this study may help other people improve their communicative effectiveness in the future.

Do I have to take part in the study?
Being in this study is your choice. If you decide to be in the study, you can change your mind and stop being in the study at any time up until the time that recordings are rated with no negative consequence. Once ratings have commenced, it will not be possible to remove your information from the results.

Can my participation in the study end early?
In addition to you being able to stop the study at any time, the study coordinator may withdraw you from this study if you do not meet the requirements for participating in the study.

Will I be paid to be in the research?
You will be reimbursed $15 for travel/parking expenses.

Will my information be kept private?
During the study we will be collecting some personal information about you. We will do everything we can to make sure that this data is kept private. None of your identifying information (e.g., name) will be published or presented publicly. All data will be stored either in a locked cabinet in a locked room or on an encrypted memory drive or computer.

The results of this study will be used in a graduate thesis, presented at academic forums, and published in an academic journal. By signing this consent form you are saying it is okay for the study staff to collect, use, and disclose information about you as described above.

After the study is done, we will still need to securely store your personal information that was collected as part of the study. At the University of Alberta, we keep data stored for at least 5 years after the end of the study. If you leave the study, we will not collect new information about you, *but we will need to keep the data that we have already collected.*

What if I have questions?
If you have any questions about the research now or later, please contact Teresa Hardy at (780) 492-7256.

The plan for this study has been reviewed for its adherence to ethical guidelines by a Research Ethics Board at the University of Alberta. If you have any questions regarding your rights as a research participant, you may contact the Health Research Ethics Board at 780-492-2615. This office has no affiliation with the study investigators.

CONSENT

Title of Study: Gender Differences in Story Retell

Principal Investigator(s): Carol Boliek, PhD **Phone Number: (780) 492-7256**
Study Coordinator: Teresa Hardy, PhD Candidate **Phone Number: (780) 492-7256**

	Yes	No
Do you understand that you have been asked to be in a research study?	☐	☐
Have you read and received a copy of the attached Information Sheet?	☐	☐
Do you understand the benefits and risks involved in taking part in this research study?	☐	☐
Have you had an opportunity to ask questions and discuss this study?	☐	☐
Do you understand that you are free to leave the study at any time, without having to give a reason?	☐	☐
Has the issue of confidentiality been explained to you?	☐	☐
Do you understand who will have access to your records, including personally identifiable information?	☐	☐
Do you understand that you will be recorded (audio and video) and these recordings will be used for future analysis?	☐	☐

Who explained this study to you? _____

I agree to take part in this study:

Signature of Research Participant

_____

_____
(Printed Name)

Date:_____

Signature of Investigator or Designee _____

Date _____

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH PARTICIPANT.

Transgender Women Communicators

PARTICIPANT CONSENT FORM

Title of Study: Talking the talk and walking the walk: Communication and quality of life in the male-to-female transgender population

Principal Investigator: Carol Boliek, PhD

Research/Study Coordinator: Teresa Hardy, Speech-Language Pathologist, PhD Candidate

---

Why am I being asked to take part in this research study?
You are being asked to be in this study because you are a person who is trans-identified and living in the female gender role and because you do not have any disorders that may impact your ability to communicate. The way a person communicates is a very important part of their identity and impacts the way others perceive their gender. We want to understand better what characteristics of communication are most important in conveying one's gender to other people, especially as it relates to being perceived as a woman. This information will help us to provide better communication feminization treatment for trans women who are undergoing gender transition.

Before you make a decision one of the researchers will go over this form with you. You are encouraged to ask questions if you feel anything needs to be made clearer. You will be given a copy of this form for your records.

What is the reason for doing the study?
We want to find out what aspects of communication make a person come across as a particular gender, what aspects of communication make a person seem masculine or feminine, and what characteristics of the voice make it sound natural/typical. This information will help identify the aspects of communication that are most important to target or change in order for an individual to be perceived as a woman by other people.

This study is being completed as part of a graduate thesis.

What will I be asked to do?
You will be asked to watch the Pink Panther cartoon, "In the Pink of the Night" and then recount the story to a researcher while being recorded. You also will be asked to fill out a series of three short questionnaires/surveys about quality of life and communication satisfaction. The visit should take 1 to 1.5 hours and you will only need to attend once. The study should take about one year to complete once all data are collected.

Information Sharing and Consent
When you arrive for your visit, a researcher will explain the study and what you will be asked to do. You then will be asked to sign a consent form to show that you agree to participate in the study, if that is the case.

Story Retell (i.e., Cartoon Description) Task
Once you arrive at the Syncrude Centre for Motion and Balance at the Glenrose Rehabilitation Hospital, you will be fitted with approximately 12 passive motion markers and a small microphone. The passive motion markers resemble small Styrofoam balls and are attached to the head, sternum, and upper body joints using adhesive pads. The microphone will be fixed to your forehead to keep it at a constant distance from your mouth. You may be provided with a standardized shirt or smock and hair band (if necessary) to allow for proper placement of the equipment. Next, you will be seated on a raised stool and you will watch the Pink Panther cartoon, "In the Pink of the Night". You will be asked to retell the events of the cartoon in as much detail as possible. Your voice will be recorded through the microphone and your movements will be recorded by a group of video cameras placed around the room, much like is done when creating computer generated imagery (CGI) movies or video games. As a result, the video recording will not show your appearance; rather, it will look like a moving stick figure. You also will be recorded repeating or reading three short phrases and holding out a vowel sound.

After the equipment is removed, you will be asked to complete a series of three short questionnaires. One asks questions about how happy you are, in general, with different aspects of your life. Another asks questions about the impact of your voice on your every day life. The last one will have you rate different aspects of your own communication in terms of how feminine and/or natural you think it is.

When you have finished the communication tasks and the questionnaires, your participation in the study will be complete.

What are the risks and discomforts?
There is a small chance that you may become fatigued while completing the study tasks. You also may feel mild awkwardness when being fitted with the recording equipment. Finally, you may become upset when completing the questionnaires, especially if you are unhappy with your quality of life or communication. You will be provided a list of trans-friendly psychologists and support groups or information about how to access SLP services to target communication feminization as needed.

There are no other known risks associated with this study; however, it is not possible to know all of the risks that may happen in a study. The researchers have taken all reasonable safeguards to minimize any known risks to a study participant.

What are the benefits to me?
You are not expected to get any benefit from being in this research study; however, this study may help other people feminize their communication as part of gender transition in the future.

Do I have to take part in the study?
Being in this study is your choice. If you decide to be in the study, you can change your mind and stop being in the study at any time up until the time that recordings are rated with no negative consequence. Once ratings have commenced, it will not be possible to remove your information from the results.

Can my participation in the study end early?
In addition to you being able to stop the study at any time, the study coordinator may withdraw you from this study if you do not meet the requirements for participating in the study.

Will I be paid to be in the research?
You will be reimbursed $15 for travel/parking expenses.

Will my information be kept private?
During the study we will be collecting some personal information about you. We will do everything we can to make sure that this data is kept private. None of your identifying information (e.g., name) will be published or presented publicly. All data will be stored either in a locked cabinet in a locked room or on an encrypted memory drive or computer.

The results of this study will be used in a graduate thesis, presented at academic forums, and published in an academic journal. By signing this consent form you are saying it is okay for the study staff to collect, use, and disclose information about you as described above.

After the study is done, we will still need to securely store your personal information that was collected as part of the study. At the University of Alberta, we keep data stored for at least 5 years after the end of the study. If you leave the study, we will not collect new information about you, *but we will need to keep the data that we have already collected.*

What if I have questions?
If you have any questions about the research now or later, please contact Teresa Hardy at (780) 492-7256.

The plan for this study has been reviewed for its adherence to ethical guidelines by a Research Ethics Board at the University of Alberta. If you have any questions regarding your rights as a research participant, you may contact the Health Research Ethics Board at 780-492-2615. This office has no affiliation with the study investigators.

CONSENT

Title of Study: Talking the talk and walking the walk: Communication and quality of life in the male-to-female transgender population

Principal Investigator(s): Carol Boliek, PhD **Phone Number: (780) 492-7256**
Study Coordinator: Teresa Hardy, PhD Candidate **Phone Number: (780) 492-7256**

	Yes	No
Do you understand that you have been asked to be in a research study?	☐	☐
Have you read and received a copy of the attached Information Sheet?	☐	☐
Do you understand the benefits and risks involved in taking part in this research study?	☐	☐
Have you had an opportunity to ask questions and discuss this study?	☐	☐
Do you understand that you are free to leave the study at any time, without having to give a reason?	☐	☐
Has the issue of confidentiality been explained to you?	☐	☐
Do you understand who will have access to your records, including personally identifiable information?	☐	☐
Do you understand that you will be recorded (audio and video) and these recordings will be used for future analysis?	☐	☐

Who explained this study to you? _____

I agree to take part in this study:

Signature of Research Participant

_____

_____
(Printed Name)

Date:_____

Signature of Investigator or Designee _____

Date _____

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH PARTICIPANT.

Raters

PARTICIPANT CONSENT FORM

Title of Study: Gender Differences in Story Retell

Principal Investigator: Carol Boliek, PhD

Research/Study Coordinator: Teresa Hardy, Speech-Language Pathologist, PhD Candidate

---

Why am I being asked to take part in this research study?
You are being asked to be in this study because you do not have special training in rating communication. The way a person communicates is a very important part of their identity and impacts the way others perceive their gender. We want to understand better what characteristics of communication are most important in conveying one's gender to other people. This information will help us to provide better treatment for people whose communication patterns do not match their gender.

Before you make a decision one of the researchers will go over this form with you. You are encouraged to ask questions if you feel anything needs to be made clearer. You will be given a copy of this form for your records.

What is the reason for doing the study?
We want to find out what aspects of communication make a person come across as a particular gender, what aspects of communication make a person seem masculine or feminine, and what characteristics of the voice make it sound natural/typical.

This study is being completed as part of a graduate thesis.

What will I be asked to do?
You will be asked to listen to and/or watch a number of samples of people describing a cartoon and to indicate whether the individual is a woman or a man. You also will be asked to rate the masculinity/femininity of the communication patterns and the naturalness of the voice. Examples of communication (e.g., voices that are very masculine, very feminine, and somewhere in between; gestures made by a man versus a woman) will be played for you to help you with your ratings. The visit should take 1.5 to 2 hours and you will only need to attend once. The study should take about one year to complete once all data are collected.

Information Sharing and Consent
When you arrive for your visit, a researcher will explain the study and what you will be asked to do. You then will be asked to sign a consent form to show that you agree to participate in the study, if that is the case.

Rating Task
You will be seated at a computer in a quiet room at the University of Alberta, the Glenrose Rehabilitation Hospital, or some other agreed-upon location. You will be given a set of headphones to wear that will cover your ears. A series of examples of communication will be played for your reference before you begin the study rating task. These examples will include: very feminine voice, very masculine voice, ambiguous voice (i.e., one that is neither very feminine nor very masculine), unnatural voice, gestures made by a woman, and gestures made by a man. You then will listen and/or watch a series of samples of different people describing a cartoon. Some of the samples will play only the audio or soundtrack, some will play only the visual (i.e., a point light display that looks like a moving stick figure) without any sound, and some will have both audio and visual. All the samples of a particular type (i.e., audio, visual, or audiovisual) will play before another type is presented. After each sample has finished playing, you will rate it using a voting box that will appear on the computer screen. When you have finished all your ratings, your participation in the study will be complete.

What are the risks and discomforts?
You may become fatigued during the rating task. There are no other known risks associated with this study; however, it is not possible to know all of the risks that may happen in a study. The researchers have taken all reasonable safeguards to minimize any known risks to a study participant.

What are the benefits to me?
You are not expected to get any benefit from being in this research study; however, this study may help other people with gender-related communication disorders in the future.

Do I have to take part in the study?
Being in this study is your choice. If you decide to be in the study, you can change your mind and stop being in the study at any time up until the time of data analysis with no negative consequence. After the time of data analysis, it will not be possible to remove your information from the results.

Can my participation in the study end early?
In addition to you being able to stop the study at any time, the study coordinator may withdraw you from this study if you do not meet the requirements for participating in the study.

Will I be paid to be in the research?
You will be reimbursed $15 for travel/parking expenses.

Will my information be kept private?
During the study we will be collecting some personal information about you. We will do everything we can to make sure that this data is kept private. None of your identifying

information (e.g., name) will be published or presented publicly. All data will be stored either in a locked cabinet in a locked room or on an encrypted memory drive or computer.

The results of this study will be used in a graduate thesis, presented at academic forums, and published in an academic journal. By signing this consent form you are saying it is okay for the study staff to collect, use, and disclose information about you as described above.

After the study is done, we will still need to securely store your personal information that was collected as part of the study. At the University of Alberta, we keep data stored for at least 5 years after the end of the study. If you leave the study, we will not collect new information about you, *but we will need to keep the data that we have already collected.*

What if I have questions?
If you have any questions about the research now or later, please contact Teresa Hardy at (780) 492-7256.

The plan for this study has been reviewed for its adherence to ethical guidelines by a Research Ethics Board at the University of Alberta. If you have any questions regarding your rights as a research participant, you may contact the Health Research Ethics Board at 780-492-2615. This office has no affiliation with the study investigators.

CONSENT

Title of Study: Gender Differences in Story Retell

Principal Investigator(s): Carol Boliek, PhD **Phone Number: (780) 492-7256**
Study Coordinator: Teresa Hardy, PhD Candidate **Phone Number: (780) 492-7256**

	Yes	No
Do you understand that you have been asked to be in a research study?	☐	☐
Have you read and received a copy of the attached Information Sheet?	☐	☐
Do you understand the benefits and risks involved in taking part in this research study?	☐	☐
Have you had an opportunity to ask questions and discuss this study?	☐	☐
Do you understand that you are free to leave the study at any time, without having to give a reason?	☐	☐
Has the issue of confidentiality been explained to you?	☐	☐
Do you understand who will have access to your records, including personally identifiable information?	☐	☐

Who explained this study to you? _____

I agree to take part in this study:

Signature of Research Participant

_____

_____
(Printed Name)

Date:_____

Signature of Investigator or Designee _____

Date _____

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH PARTICIPANT

APPENDIX E – DEMOGRAPHIC INFORMATION DATA COLLECTION FORMS

Cisgender Communicators

Participant Number: _____

Date: _____

INCLUSION/EXCLUSION CRITERIA	
Fluent English speaker	
Full use of upper body	
No neurogenic communication disorder	
VARIABLE	**DATA**
Age in years	
Cisgender man or woman?	
Ethnic Origin British Isles French Eastern European Aboriginal Other North American Caribbean Latin, Central, and South American African Arab West Asian South Asian East and Southeast Asian Oceania	
Smoking – past (yes/no)	
Smoking – time of assessment (yes/no)	

Transgender Women Communicators

Participant Number: _____

Date: _____

INCLUSION/EXCLUSION CRITERIA	
Fluent English speaker	
Full use of upper body	
No neurogenic communication disorder	
Living in female role majority of time	
Living in female role at least 6 months	
VARIABLE	**DATA**
Demographic Information	
Age in years	
Time spent living as female for a percentage of the day (years)	
Time spent living full-time as a female (years)	
Smoking – past (yes/no)	
Smoking – time of study (yes/no)	
Receiving hormone therapy (yes/no)	
Duration of communication feminization treatment (months)	
Time since treatment (months)	
Reassignment surgery (yes/no)	

Ethnic Origin British Isles French Eastern European Aboriginal Other North American Caribbean Latin, Central, and South American African Arab West Asian South Asian East and Southeast Asian Oceania	
Self-Perception/QOL/Comm. Satisfaction	
QLI Overall QoL (Total)	
TVQ^{MtF} (Total score)	
Perceived Vocal Femininity	
Perceived Gestural Femininity	
Perceived Communication Femininity	
Perceived Voice Satisfaction	
Perceived Communication Satisfaction	

Raters

Participant Number: _____

Date: _____

INCLUSION/EXCLUSION CRITERIA	
Proficient in English	
Pass hearing screen	
No uncorrected visual deficits	
No training in communication disorders	
No language/learning/cognitive disabilities	
VARIABLE	**DATA**
Age in years	
Sex (M/F/Other – specify)	
Ethnic Origin British Isles French Aboriginal Other North American Caribbean Latin, Central, and South American African Arab West Asian South Asian East and Southeast Asian Oceania	
Highest level of education	

APPENDIX F – SELF-RATINGS FORM (TRANSGENDER WOMEN)

Personal Perceptions of Communication

Please indicate how you feel about different aspects of your communication using the rating scales provided below. Circle the value that best reflects your opinion of your own communication right now. Be sure to choose a number rather than the point in between.

1. My voice sounds

1	2	3	4	5	6	7
Not at all Feminine			Somewhat Feminine			Very Feminine

2. My gestures look

1	2	3	4	5	6	7
Not at all Feminine			Somewhat Feminine			Very Feminine

3. Overall, my communication is

1	2	3	4	5	6	7
Not at all Feminine			Somewhat Feminine			Very Feminine

4. I am _____ with the way my voice sounds.

1	2	3	4	5	6	7
Not at all Satisfied			Somewhat Satisfied			Very Satisfied

5. I am _____ with my overall communication.

1	2	3	4	5	6	7
Not at all Satisfied			Somewhat Satisfied		Satisfied	Very

APPENDIX G – POST-RATING ALLY STATUS AND SEXUAL ORIENTATION QUESTIONNAIRE

Participant Number: _____

Date: _____

1. How many of your family, friends and/or close colleagues identify with the greater lesbian, gay, bisexual, transgender, and queer (LGBTQ) community? (Note: this may include sexual orientations or gender identities not encompassed by the LGBTQ acronym).

Please circle your response.

None (0%) Few (<20%) Many (20% - 50%) Most (>50%)

2. How would your describe your own sexual orientation?

1	2	3	4	5	6	7	8	9
Very Heterosexual								Very Homosexual