2010

The effects of voice pitch and resonances on assessments of speaker size, masculinity, and attractiveness

Pisanski, Katarzyna Alicja

Lethbridge, Alta. : University of Lethbridge, Dept. of Psychology, 2010

http://hdl.handle.net/10133/2549

THE EFFECTS OF VOICE PITCH AND RESONANCES ON ASSESSMENTS OF SPEAKER SIZE, MASCULINITY, AND ATTRACTIVENESS

Katarzyna Alicja Pisanski
B.A., McMaster University, 2007

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

MASTER OF SCIENCE

Department of Psychology
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

ABSTRACT

The human voice might have been shaped by sexual selection. Hence, voice fundamental (F_0, or pitch) and formant frequencies (Fn, or timbre) are proposed to convey fitness cues germane to rivals and potential mates. First, I confirm the independent effects of F_0 and Fn on listeners' assessments of speaker size, masculinity, and attractiveness. Second, I quantify the just-noticeable differences in both vocal features and then place F_0 and Fn cues in conflict by equally discriminable amounts to test their relative influence on such voice-based social judgments. Results revealed a greater relative role of Fn in listeners' ratings of all three dimensions, suggesting that these dimensions might all be cued more reliably by Fn than F_0. Alternatively, given post-hoc principal component analyses that revealed considerable overlap in ratings of size, masculinity, and attractiveness, listeners' conceptions of these dimensions may not be independent despite a research tradition that assumes they are.

ACKNOWLEDGMENTS

First and foremost, I extend my sincere gratitude and appreciation to my supervisor, Drew Rendall, for his unwavering support and limitless contributions to my master's thesis and research. Throughout these past two years, Drew has been my guide, my teacher, and my mentor. Our philosophical conversations have taken my mind to new places. His critiques and advice encourage humility and have helped me to become a better academic, while his understanding and support have given me the strength to reach my goals. At the same time, Drew has treated me as an equal, as an academic colleague, and for this autonomy and respect I am sincerely humbled. I cannot begin to measure the impact that Drew has had, and surely will continue to have, on my research; nor can I find the words to express to him how truly indebted I am for the lessons, skills, and experience I leave here with.

Likewise, I am appreciative of the counsel and encouragement that I have received from my committee members, John Vokey and Louise Barrett. In addition to regular feedback on experimental design, methodology, theory, and statistics, John and Louise have encouraged my involvement in local conferences and provided me with invaluable letters of reference. Perhaps equally as important, I have thoroughly enjoyed our countless, informal exchanges; their freethinking brilliance, originality, and general optimism toward life have inspired me on both a scholarly and personal level. I would like to thank John and Louise not only for their mentorship but also for their genuine companionship.

I also offer my thanks to all of the remaining faculty, students and staff in the Department of Psychology and neighbouring departments for their support. In particular, I would like to thank Martin Lalumière, Scott Allen, Peter Henzi, Paul Vasey, Fang Fang Li, Margaret Forgie, Danny Krupp, and Sandeep Mishra for their guidance and for any contributions each has made to my research program. I would also like to thank my lab members past and present, Shannon Digweed, Brandon Yardy, and Alan Nielsen, for their assistance, advice, and of course, comradery and fellowship. Lastly, I thank Leanne Wehlage-Ellis for the immense amount of administrative work that she does each day in order to keep the department (and its perhaps less organized constituents) functioning efficiently.

I extend a special thanks to a fellow graduate, my partner and best friend, Mitchell LaPointe, who for two years has offered me endless intellectual stimulation and emotional support. Mitch, thank you first for all of your direct contributions including theoretical advice and edits. More importantly, thank you for your companionship, your understanding, and for constantly motivating me to be my very best.

I also thank my wonderful family who, even from such a distance, have been a constant guiding light in my life. To my father and mother, Maciej and Ewa, thank you for the values you have instilled in me, including the pursuit of knowledge and a quest for goodness. Thank you for the unconditional love that has made it possible for me to dream big and the emotional support that has allowed me to reach those dreams, one step at a time. To my big brother, Piotr, you are my greatest mentor of all. Your love for learning is an inspiration to everyone who meets you and your brilliance is something to aspire to.

Thank you for your many contributions to my work, both technical and conjectural, and for always having been my biggest fan.

Finally, I give thanks to the Natural Sciences and Engineering Research Council of Canada for Discovery Grants for Drew Rendall as well as the Government of Alberta, the Department of Psychology, the School of Graduate Studies, and the Graduate Students' Association for directly funding my research and research-related travel.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BMI	Body-mass index
bVt	B(vowel)t
dB	Decibel
D_f	Formant dispersion
f	Female
FA	Fluctuating asymmetry
fMRI	Functional magnetic resonance imaging
F_0	Fundamental frequency (pitch)
Fn	Formant or resonance frequency (timbre)
HD	High definition
Hz	Hertz
JND	Just-noticeable difference
m	Male
ms	Milliseconds
PCA	Principal component analysis
pt	Point
RT	Response time
s	Second
SHR	Shoulder-hip ratio
spl	Sound pressure level
v.	Version
VTL	Vocal-tract length
WHR	Waist-hip ratio

CHAPTER ONE

INTRODUCTION

1.1 OVERVIEW

Vertebrate vocalizations, such as the roars of male red deer (Clutton-Brock & Albon, 1979) the calls of monkeys (Cowlishaw, 1996; Harris, Fitch, Goldstein & Fashing, 2006; Rendall, Kollias, Ney, & Lloyd, 2005) the songs of birds (Eriksson & Wallin, 1986), and the speech of humans can serve a variety of functions. One basic function they can serve is to cue important indexical characteristics about vocalizers, such as their age, sex, or individual identity, that might be relevant in mediating many kinds of routine behavioural and social interactions with other individuals. It is also likely that vocalizations have evolved, in part, under the pressures of sexual selection (Darwin, 1871), and so function to advertise additional characteristics of a vocalizer that might be important in the context of competing for and attracting mates.

For example, the quality of song in male barn swallows, *Hirundo rustica,* reflects the androgen levels, body mass, and body condition of the vocalizer that are characteristics that are likely to affect a male’s competitive ability and potential attractiveness to female mates (Galeotti, Saino, Sacchi & Møller, 1997)**.** Indeed, different features of male bird song have been demonstrated to attract female mates in natural environments (Eriksson & Wallin, 1986).

Research on human vocal signaling has focused primarily on speech and its role in communicating linguistic information. It is possible that the human voice has likewise been shaped by sexual selection, and, in fact, voice-based relationships like those documented in animals have also been reported in humans. For example, in human males, deep voices have been shown to correlate with high testosterone levels and height (Bruckert, Liénard, Lacroix, Kreutzer, & Leboucher, 2006) and are preferred by females to male voices of higher frequencies (Bruckert *et al.*, 2006; Collins, 2000; Feinberg, Jones, Little, Burt & Perrett, 2005b; Saxton, Caryl, & Roberts, 2006).

In order to understand how the human voice might have been modified by sexual selection, research has often focused on distinguishing perceptual elements of the voice that might plausibly be associated with specific characteristics of vocalizers that would be relevant in mate competition and mate-choice. In this thesis, I report a series of experiments focusing specifically on the potential role of voice pitch (F_0) and formant frequencies (*F*n) in cueing socially relevant speaker characteristics related to overall body-size, masculinity or femininity, and attractiveness. A goal of the work is to test alternative hypotheses that stress the differential salience of these two voice features in listeners' assessments of speakers and to try to converge on viable social and functional explanations for the outcomes.

1.2 THEORETICAL BACKGROUND

1.2.1 Vertebrate Vocal Production

In most vertebrates, including mammals, birds, amphibians, and reptiles, vocalizations are produced by an acoustic source (e.g., the larynx) that acts to convert air flow from the lungs into acoustic energy. Typically, air produced by the lungs oscillates the vocal folds within the larynx and is then channeled through the vocal-tract and mouth cavity before exiting the nostrils and lips as vocal output or phonation (Figure 1.1). The supralaryngeal vocal-tract thus acts to filter the speech produced by the sound source (Fitch & Hauser, 2003).

In humans, the two most perceptually salient elements of speech, pitch and 'timbre' (Bruckert *et al.,* 2006; Baumann & Belin, 2010), are based on the fundamental and formant frequencies of the voice, respectively. The fundamental frequency (F_0), associated with the percept of pitch, is determined by the rate of vocal fold vibration. The rate at which the vocal folds vibrate is contingent on the shape and size of the folds. Thus, the average adult male's pitch is 100 Hertz (Hz) whereas adult females, naturally possessing smaller and shorter vocal folds, speak with an average pitch of 200 Hz. Young children have even shorter vocal folds and speak with an average pitch that is higher still at 400-500 Hz (Lieberman & Blumstein, 1988).

A second perceptually salient dimension of the human voice, roughly akin to the percept of timbre, is a direct product of vocal-tract resonances termed *formant frequencies*. Formant frequencies (Fn) represent bands of energy that have been permitted to pass through an individual's supralaryngeal airway. Consequently, formant frequencies are largely determined by the length and shape of the vocal-tract and mouth cavity that act to filter all speech sounds (Lieberman & Blumstein, 1988). They have been shown to correlate negatively with vocal-tract length in monkeys, humans, and some other mammals and birds (Fitch 1997; Fitch, 1999; Fitch & Giedd, 1999; Reide & Fitch, 1999). Formants are denoted from lowest to highest in frequency beginning with

$F1$ through to as many as $F7$, though it is the first four formants that are the strongest and likely the most salient perceptually.

In accordance with the source-filter theory of speech production (Müller, 1848), the frequencies produced by the vocal source, the larynx (i.e., the fundamental frequency and its harmonics), and those produced by the filtering of the vocal-tract (i.e., formants) are independent of one another. Whereas F_0 is determined solely by the rate at which the soft vocal folds vibrate, formants are shaped by the hard-tissue of the vocal-tract. Hard-tissue is less malleable than soft-tissue; as a consequence, larger individuals with longer vocal-tracts ought to have lower formant frequencies but not necessarily a lower fundamental (Fitch, 2000).

Today's computer technology allows researchers to readily transform sound recordings into visual representations. A speech sound can be charted as a wave-form representing variations in air-pressure or, by Fourier analysis, as a spectrum on a spectrogram with frequency in Hz on the Y-axis and time on the X-axis (Baken, 1987). Figure 1.2 shows a spectrogram of the word "butt" spoken by an adult human male. The darkness or lightness of the bands on the spectrogram represents the relative strength or amplitude of the sound at a given frequency (Lieberman & Blumstein, 1988). The fundamental frequency of the sound is represented by the lowest band of energy on the spectrogram. The formants are represented as the broad regions of energy emphasis above F_0, beginning with the lowest formant, $F1$, at roughly 400 Hz and climbing in frequency to the highest visible formant, $F3$, at roughly 2500 Hz.

1.2.2 Indexical Voice Cues

The voice carries cues to a variety of personal dimensions of a speaker such as their age (Bruckert *et al.*, 2006; Collins & Missing, 2003; Evans, Neave, & Wakelin, 2006; Feinberg, 2004), sex (Childers & Wu, 1991; Jenkins, 1998; Owren, Berkowitz, & Bachorowski, 2007), and identity (Owren & Cardillo, 2006; Rendall, Notman, & Owren, 2009). Sex or gender cues are especially salient. The human vocal folds are sexually dimorphic; increased exposure to testosterone during puberty permanently enlarges and lengthens the male vocal folds by an average of 63% as compared to only 34% in females (Jenkins, 1998). This developmental effect slows the rate at which the fully developed vocal folds of adult males vibrate, creating a lower F_0 and the associated perception of a lower-pitched voice in men. This vocal cue to an individual's sex is reliable and shows high consensus among listeners' ratings of speaker-sex from voices alone (Childers & Wu, 1991).

In addition to these marked and predictable frequency differences between male and female voices, there is also appreciable within-sex variation in F_0 and formants. The F_0 of adult human males, for example, may range anywhere from 80 to 300 Hz (Lieberman & Blumstein, 1988). Such individual variation provides a basis for recognizing specific individuals (Owren & Cardillo, 2006) and affects the types of personal characteristics we attribute to them (e.g., how masculine or attractive we perceive them to be; Bruckert *et al.*, 2006; Feinberg, Jones, Little, Burt, & Perrett, 2005b).

1.2.3 Sexual Selection of the Human Voice

Vertebrate vocalizations may have been under selective pressure from both intersexual selection (female mate-choice) and intrasexual selection (male-male competition; for review see Puts, 2010). A competitive advantage of, and a female preference for, deeper vocal frequencies in males may help in part to explain the large dimorphism in human male and female F_0.

Moreover, if the human voice has been sexually selected, F_0 and Fn could provide cues to an individual's general health or relative fitness; research suggests that, in fact, they do. Cues to immediate health, such as whether a person smokes or has a cold, are salient in the voice through audible perturbations of F_0 (Feinberg, 2004). Cues to more stable predictors of health, such as fluctuating asymmetry (FA), may also be traceable through the voice. Low FA is a sign of developmental stability and has been found to correlate with number of sexual partners (Thornhill & Gangestad, 1994) and, more recently, with voice quality (Hughes, Harrison, & Gallup, 2002).

Perhaps most important, F_0 and Fn offer clues to underlying levels of circulating hormones: low F_0 or Fn can predict high levels of testosterone in men (Bruckert *et al.*, 2006; Dabbs & Mallinger, 1999) and high F_0 can predict high levels of estrogen in women (Abitbol, Abitbol & Abitbol, 1999). Accordingly, research shows that high F_0 can act as a reliable cue to ovulation in women (Bryant & Haselton, 2009). It may also predict health and fitness because estrogen correlates positively with female fertility and developmental health (Alonso & Rosenfield, 2002). Testosterone, that compromises immunocompetence, can predict resistance to lowered immunity and high developmental stability in healthy males (Folstad & Karter, 1992). Consequently, deep voices in men and high voices in women might reliably cue important fitness-related dimensions of speakers.

Apart from cueing age and sex, the voice can cue a host of other personal characteristics relevant especially to social exchange and mate selection, such as cues to health and reproductive value (fitness). As perceivers, we might interpret cues to relative fitness in a variety of ways. For example, the voice can affect how attractive, dominant, masculine, or even large we judge an individual to be. Attractiveness, masculinity and body-size are relevant qualities when choosing a mate or evaluating a rival because they tend to co-vary with other dimensions of health, fitness and strength. Vocal attractiveness and vocal masculinity, for example, have been found to predict an individual's number of sex partners (Hughes, Dispenza, & Gallup, 2004; Puts, 2005) and number of children (Apicella, Feinberg, & Marlowe, 2007).

1.2.4 Vocal Cues of Attractiveness

Let us consider facial attractiveness in humans as a case in point. In general, the faces of men with high testosterone levels and those of women with high estrogen levels are rated as more attractive than individuals with low levels of either hormone (Feinberg, 2008) cross-culturally (Rhodes, 2005). What appears to be attractive about these individuals are precisely those facial features that are sculpted by steroid hormones, such as face shape (Feinberg, 2008), and those related to developmental stability, such as facial symmetry (Rhodes, Yoshikawa, Clark, Lee, McKay, & Akamatsu, 2001; Thornhill & Gangestad 1999).

Vocal attractiveness, like facial attractiveness, appears to be related to underlying hormone levels and developmental stability. As one may predict by the logic of sexual selection, females rate low-pitched male voices as more attractive than high-pitched male voices (Bruckert *et al.*, 2006; Collins, 2000; Feinberg *et al.*, 2005b; Saxton *et al.*, 2006). This effect is especially salient when women are ovulating (Feinberg, Jones, Law Smith, Moore, DeBruine, Cornwell, Hillier, & Perrett, 2006; Puts, 2005), which is when "good genes" in a mate are predicted to be particularly important in their mate-choice (Feinberg, 2008; Hughes & Gallup, 2008). Conversely, males prefer high-pitched female voices over low-pitched female voices (Collins & Missing, 2003; Feinberg, DeBruine, Benedict, & Perrett, 2008a).

In addition, vocal attractiveness has been found to correlate with attributions of kindness and honesty in women and assertiveness and dominance in men (Berry, 1990). Such attributions are reminiscent of a generally fallible human bias to associate attractive physical traits, such as an attractive voice, with positive personality traits (Zuckerman & Driver, 1989). Nevertheless, this "vocal attractiveness stereotype" may not be entirely unwarranted. Attractive voices may, after all, belong to relatively healthier and more reproductively and socially successful people (Apicella *et al.*, 2007).

1.2.5 Vocal Cues of Dominance and Masculinity

The voice acts to cue attractiveness to potential mates; however, it may also act to cue dominance and masculinity to potential rivals. Intrasexual male competition over resources and mates can create selection pressure on male vocalizations.

Male gibbons (*Hylobates*) and red deer stags (*Cervus elaphus*) are among many species in which males defend territories, guard mates, and successfully establish dominance through loud vocalizations (Cowlishaw, 1996; Fitch & Hauser, 2003). Acoustic displays of aggression are much safer than physical displays and can be equally effective when such vocal cues are salient and reliable. In male gibbons, vocalizations are reduced when energy constraints, such as a lack of food, are introduced (Cowlishaw, 1996), suggesting that routine vocalizing is costly because it requires energy. In male red deer, both roaring and fighting co-vary positively with physical condition and resilience (Clutton-Brock & Albon, 1979). Male vocalizations, therefore, may act as honest cues to a high energy level in gibbons and physical prowess in red deer.

In human males, F_0 tends to correlate negatively with circulating testosterone and *F*n tends to correlate negatively with height (for cues of body-size see Section 1.2.6). Both vocal features may, therefore, reliably cue masculinity and dominance. Accordingly, low F_0 and *F*n elicit high ratings of social and physical dominance (Puts, Hodges, Cardenas, & Gaulin, 2007) as well as high ratings of masculinity in men (Feinberg *et al.*, 2005b; Feinberg, 2008). Collins (2000) found that men with lower F_0 were also judged to have hairier chests and larger muscles than men with higher F_0, though these listener perceptions did not align with any true relationship between hairiness, muscularity, and voice pitch in this sample of men. In a separate study, listeners were able to accurately assess a male speaker’s upper-body-strength from voice recordings alone (Sell, Bryant, Cosmides, Tooby, Sznycer, von Rueden, Krauss, & Gurven, 2010).

In women, high F_0 may correlate with levels of estrogen (Abitbol *et al.*, 1999). Consequently, tonal features of the voice may change across a woman's menstrual cycle, during pregnancy, or after menopause (Hughes & Gallup, 2008). These tonal changes have been shown to cue ovulation (Bryant & Haselton, 2009). Related research regarding the effects of F_0 and formant frequencies on perceptions of femininity in female voices is minimal compared to research on male voices. Nonetheless, preliminary studies, including those reported here, suggest that female voices with higher F_0 (and perhaps higher estrogen levels) are rated as more feminine than their deeper-voiced counterparts (Feinberg *et al.*, 2008a).

1.2.6 Vocal Cues of Body-Size

Body-size is another important factor in mate-choice and in male-male competition. In many primate species, including humans, larger bodied males fare better in both domains (Alexander, Hoogland, Howard, Noonan, & Sherman, 1979; Ghazanfar & Rendall, 2008; Pawlowski, Dunbar & Lipowicz, 2000). Interestingly, in the same way that the vocalizations of red deer stags provide information about relative strength and dominance, vertebrate vocalization may also function to provide information about body-size. In colobus monkeys, for example, larger bodied individuals have been shown to vocalize at lower frequencies than smaller individuals (Harris *et al.*, 2006). A similar negative relationship between body-size and vocal frequencies has been found in many bird species (Fitch, 1999), including the tawny owl (Appleby & Redpath, 1997) and barn swallow (Galeotti *et al.*, 1997).

In humans, there is a clear relationship between vocal frequency and *perceived* body-size in which speakers with deep voices are consistently rated as large, while speakers with relatively high voices are rated as smaller across both sexes (Collins, 2000; Feinberg *et al.*, 2005b). Despite this consensus in people's assessments of body-size (González, 2006), three questions remain to be answered. First, do vocal characteristics actually correlate with real height or weight in humans? Second, are people's assessments of body-size veridical? Third, which characteristics of the voice are driving the assessments and why?

Adult humans are, of course, physically larger than prepubescent children. In view of that, pre-pubescent children speak with an average pitch of 500 Hz and adults with an average pitch of 150 Hz (Lieberman & Blumstein, 1988). Humans, like most other species, further exhibit sexual dimorphism in adult body-size. Western adult males are, on average, 8% taller and 17% heavier than adult females (Ogden, Fryar, Carroll, & Flegal, 2004). Of course, this relatively small difference in body-size between the adult sexes is not sufficient to explain the two-fold difference between male and female baseline F_0. Nevertheless, frequency changes in the voice accurately cue body-size if one is comparing children and adults or men and women (Rendall *et al.*, 2005). Whether F_0 or Fn can accurately cue body-size *within* adult sex is debatable.

While previous studies have failed to find a correlation between F_0 and *height* in particular (González, 2004; Künzel, 1989; van Dommelen & Moxness, 1995), some have found support for a negative correlation between F_0 and other measures of body-size in humans. For instance, Evans and her colleagues (2006) found that low F_0 predicted heaviness in men. Collins & Missing (2003) found the same negative relationship between F_0 and weight in women. In contrast, Fn has more consistently been found to correlate with several measures of size in humans. Low Fn can predict greater height (Collins & Missing, 2003; González, 2004; Greisbach, 1999) and weight (González, 2004) in women, and greater height (Evans *et al.*, 2006; Greisbach, 1999; Rendall *et al.*, 2005; Rendall, Vokey, & Nemeth, 2007) and weight (Evans *et al.*, 2006; González, 2004) in men. A handful of studies have, nonetheless, failed to replicate correlations between formants and body-size (Bruckert *et al.*, 2006; van Dommelen & Moxness, 1995).

In support of demonstrating a real relationship between the voice and the body, a small number of studies have also tested for correlations between the voice and body-shape. Body-shape, like size, might also provide important cues to an individual's health and reproductive value (Singh, 1993). High shoulder-hip ratio (SHR) in men and low waist-hip ratio (WHR) in women, for example, can sometimes predict high levels of testosterone and estrogen, respectively. These specific body-shape configurations have also been found to predict attractiveness (Singh, 1993), age of first intercourse, and number of sexual partners (Hughes *et al.*, 2004). Appropriately, low F_0 and low formant dispersion (D_f, the average difference in frequency between successive formants, Fitch 1997) appear to predict a larger SHR in men (Evans *et al.*, 2006). F_0 has also been found to correlate negatively with shoulder and chest circumference in men (Evans *et al.*, 2006) but not with neck, skull, waist or hip circumference. Likewise, D_f has been found to correlate negatively with men's shoulder, chest, neck, and waist circumference but not with skull and hip circumference (Evans *et al.*, 2006). Comparable studies have not yet been done for WHR and vocal features in women.

In summary, studies show that F_0 and Fn (or D_f) can, in some cases but not in others, accurately predict height, weight or body configuration. Although perceptual studies show that listeners, in all cases, rate persons with deeper voices as larger and taller, few studies have tested the relationship between real body-size and people's perceptions of body-size from the voice. Such studies require the use of real, rather than synthetic, voice stimuli (González, 2006). The few studies that have done so, noted below, have demonstrated that the accuracy of people's assessments is in fact as variable as the genuine relationships between size and the voice. In other words, sometimes listeners are correct in their voice-based body-size assessments (Bruckert *et al.*, 2006; Collins, 2000; Dusan & Deng, 1999; González & Oliver, 2004; Ives, Smith & Patterson, 2005; Rendall *et al.*, 2005, 2007; Smith, Patterson, Turner, Kawahara, & Irino, 2005) and sometimes they are not (Bruckert *et al.*, 2006; González, 2003; van Dommelen & Moxness, 1995). These inconsistent findings might in part reflect different methodologies across studies. On the other hand, the accuracy of body-size assessments may depend on which elements of the voice (e.g., F_0 and/or Fn) listeners are attending to.

1.2.6.1 Vocal Cues of Body-Size: Formants vs. Pitch

According to Fitch (1994, 1997, 2000; Fitch & Hauser, 2003), formants are more reliable than F_0 as cues to height in mammals. This is because, although F_0 depends on the size and subsequent vibration of the vocal folds, mammalian formant frequencies are largely dependent on vocal-tract length (VTL). The hard-tissues of the supralaryngeal vocal-tract are far less malleable than the soft vocal folds. Because the vocal folds are constructed of soft-tissue, their development is not constrained by skeletal architecture while VTL is

(Fitch, 2000). As a result, Fitch claims, taller individuals *necessarily* have longer vocal-tracts and lower formant frequencies but may not necessarily have longer vocal folds and a low-pitched voice, and this should hold both within and across sexes.

Indeed, a number of recent mammalian studies have reported a strong, three-way relationship between VTL, body-size and *F*n in rhesus macaques (Fitch, 2007), domestic dogs (Riede & Fitch, 1999), and humans (Fitch & Giedd, 1999; Dusan & Deng, 1999; Dusan, 2005). However, because these particular studies collapse across age and sex, the correlations reported in them are likely to exaggerate the true relationship between VTL and *F*n *within* age and sex categories (for discussion see Rendall *et al.,* 2005). And, because what is of interest to us is how the voice has been sexually selected through female mate-choice and male-male competition, these cross-category comparisons are of little use.

Nevertheless, while evidence for a *three*-way relationship between VTL, size and *F*n within the adult sexes is lacking, there is at least good evidence for a *two*-way relationship between body-size and *F*n within adult males of several species, including, red deer (Reby & McComb, 2003), colobus monkeys (Harris *et al.,* 2006) and giant pandas (Charlton *et al.*, 2009), where larger bodied males have lower formant frequencies than smaller males. This two-way relationship between size and *F*n has also been demonstrated in adult men and women (see Table 2.1).

Fitch's theory allows us to make some predictions about listener's disparate sensitivity to pitch and formant cues in assessments of size. Thus, as a result of the specific anatomical (VTL or size) constraints on *F*n and not F_0, we might predict that listeners will attend only to the former when attempting to assess height. While it is clear that listeners *do* attend to *F*n when making body-size assessments (Collins & Missing, 2003; Feinberg *et al.*, 2005b; von Dommelen & Moxness, 1995), paradoxically, they also attend to F_0 (Collins, 2000; Feinberg *et al.,* 2005b; van Dommelen & Moxness, 1995). Thus, the question remains why listeners attend to F_0 given that it does not reliably correlate with height. Possible explanations for this misattribution bias will be addressed in subsequent sections of this thesis.

1.2.7 Predicted Role of Pitch and Formants in Listeners' Assessments

In summary, males of many vertebrate species, especially mammals, emit low-frequency vocalizations that may act to directly cue social status and body-size while indirectly cueing health and mate-quality. Male vocalizations may function as evolved, secondary sex characteristics. Recent human research suggests that the female voice may have also evolved to cue indexical information, though the mechanisms and roles of female vocalizations are poorly understood.

In humans, listeners use multiple features of the voice as cues to a speaker's body-size, masculinity or femininity, and attractiveness, among other traits. Table 1.1 summarizes results from a series of studies examining the independent effects of F_0 and *F*n on listener's *perceptions* of these three speaker dimensions. In brief, the literature suggests that low F_0 and *F*n independently cue greater body-size and masculinity in both male and female speakers. In addition, low F_0 and *F*n are attractive in male speakers while high F_0 and *F*n are attractive in female speakers.

Table 1.2 summarizes a series of studies examining the *real,* physical and physiological relationships between vocal features and measures of size and masculinity/femininity. Although voice-based *assessments* of biosocial dimensions enjoy high consensus among raters, the appropriateness and accuracy of these assessments is uncertain. It also remains uncertain which elements of the voice are driving vocal assessments and precisely how body-size, masculinity, and attractiveness relate to one another conceptually. Thus, testing the relationships between F_0 and Fn in both sexes, and how they relate to various possible biosocial dimensions relevant in mate competition and mate-choice, is a critical prerequisite to a fuller understanding of how and to what degree the human voice has been modified by sexual selection.

1.3 THESIS OUTLINE

The aim of this thesis is to gain a better and more comprehensive understanding of the effects of F_0 and Fn on listeners' perceptions of speaker size, masculinity and attractiveness. In Chapter Two, I report findings from an experiment designed to test the independent and joint (combined and unidirectional) effects of F_0 and Fn on listeners' voice-based assessments. I further test whether listeners' perceptions change as a result of rating computer-manipulated versus natural voices or as a function of speaker or listener sex.

Chapter Three attempts to establish independent discrimination thresholds for F_0 and Fn. I then report on the relative contributions of F_0 versus Fn in listener's assessments of size, masculinity and attractiveness, when the differences in these features are controlled and made equally discriminable to listeners. I predict that, because of its strong association with steroid hormones, listeners will track F_0 over Fn in assessments of masculinity or femininity and attractiveness. I further predict that, because of its association with VTL and height, listeners will track Fn over F_0 in assessments of body-size.

Chapter Four explores inter-relationships among listeners' ratings of size, masculinity and attractiveness, that appear not to be independent of one another. I offer an account of the observed inter-relationships that proposes that the overlap in ratings of different dimensions reflects overlap in listeners' conceptions of the dimensions.

Chapter Five provides a summary of findings and their significance. It also discusses some limitations of the work reported and deals with some unresolved issues and directions for future research.

Table 1.1 The predictive value of speakers' F_0 and Fn in listeners' assessments of size, masculinity or femininity, and attractiveness. For each rating dimension, the direction of a significant effect with respect to F_0 and Fn is indicated with arrows (∧ or ∨) and, in each case, the voice condition correlates with ratings of larger size, greater masculinity (or femininity) and higher attractiveness. *ns* denotes nonsignificant effects ($P > 0.05$). A blank cell indicates that a relationship was not tested.

Author(s)	**Year**	**Male Speakers**					
		body-size		**masculinity**		**attractiveness**	
		F_0	Fn	F_0	Fn	F_0	Fn
Puts *et al.*	**2007**			∨	∨		
Puts *et al.*	**2006**			∨			
Saxton *et al.*	**2006**					∨	
Feinberg *et al.*	**2005b**	∨	∨	∨	∨	∨	*ns*
Puts	**2005**					∨	
Collins	**2000**	∨		∨		∨	
van Dommelen & Moxness	**1995**	∨	∨				
Author(s)	**Year**	**Female Speakers**					
		body-size		**femininity**		**attractiveness**	
		F_0	Fn	F_0	Fn	F_0	Fn
Collins & Missing	**2003**					∧	∧
van Dommelen & Moxness	**1995**	∨	∨				
Feinberg *et al.*	**2008a**					∧	

Table 1.2 Physical relationships between F_0, *F*n, body-size, and masculinity/femininity in men and women. For each physical dimension, the direction of a significant effect with respect to F_0 and *F*n is indicated with arrows (∧ or ∨) and, in each case, the voice condition correlates with measures of larger size, greater masculinity (or femininity) and higher attractiveness. *ns* denotes nonsignificant effects ($P > 0.05$). A blank cell indicates that a relationship was not tested.

Author(s)	**Year**	**MEN**					
		height		**weight**		**masculinity**	
		***F*0**	***F*n**	***F*0**	***F*n**	***F*0**	***F*n**
Rendall *et al.*	**2007**	∨[a]	∨				
Bruckert *et al.*	**2006**		*ns*			*ns*	∨
Evans *et al.*	**2006**	*ns*	∨	∨	∨		
Rendall *et al.*	**2005**	*ns*	∨	*ns*	*ns*		
González	**2004**	*ns*	∨	*ns*	∨		
Dabbs & Mallinger	**1999**					∨	
Griesbach	**1999**		∨				
van Dommelen & Moxness	**1995**	*ns*	*ns*	*ns*	*ns*		
Künzel	**1989**	*ns*		*ns*			
Author(s)	**Year**	**WOMEN**					
		height		**weight**		**femininity**	
		***F*0**	***F*n**	***F*0**	***F*n**	***F*0**	***F*n**
Feinberg *et al.*	**2005a**					∧	
Rendall *et al.*	**2005**	*ns*	*ns*	*ns*	*ns*		
González	**2004**	*ns*	∨	*ns*	∨		
Collins & Missing	**2003**	*ns*	∨	∨	*ns*		
Abitol *et al.*	**1999**					∧	
Griesbach	**1999**		∨				
van Dommelen & Moxness	**1995**	*ns*	*ns*	*ns*	*ns*		
Künzel	**1989**	*ns*		*ns*			

a. F_0 was predictive of height differences only in comparisons of very short and very tall men with sizeable height disparities (14-27 cm).

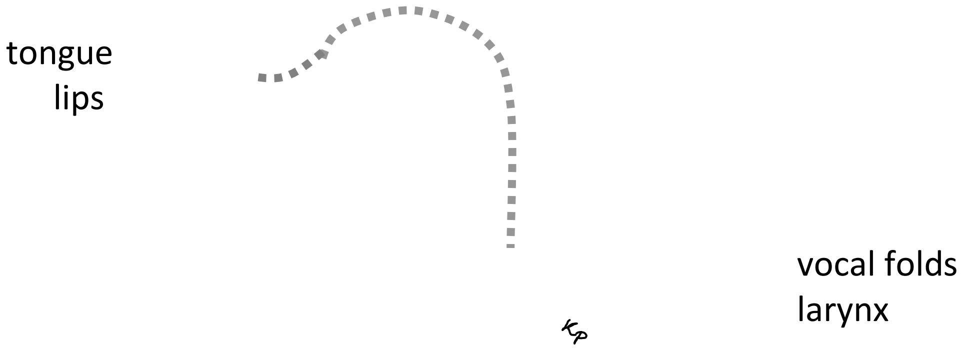


Figure 1.1. A labeled diagram of the human vocal-tract. (Diagram was drawn by author and adapted in part from Fitch, 1994).

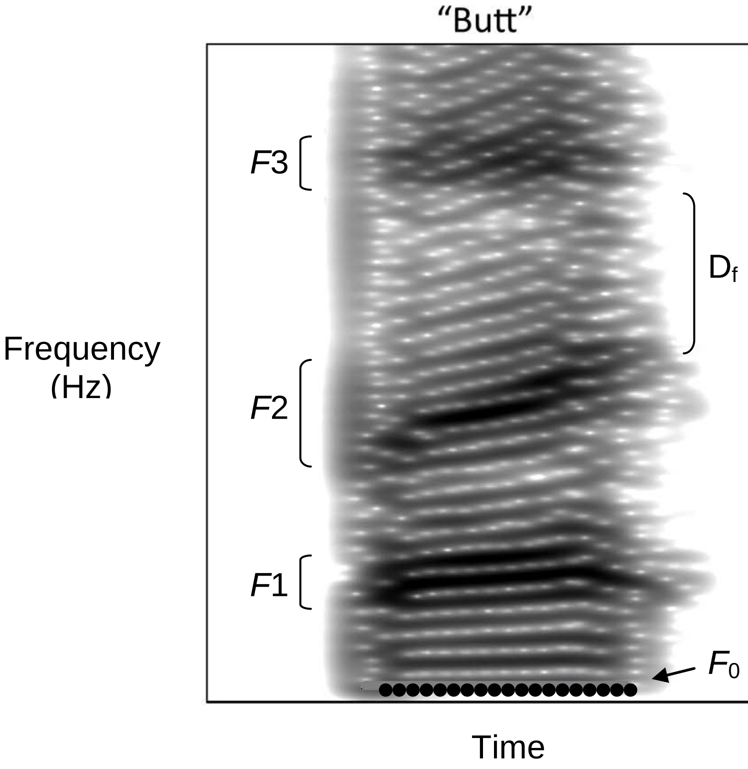


Figure 1.2. A labeled spectrogram of the word "butt" spoken by an adult human male. The darkness or lightness of the bands on the spectrogram represents the relative amplitude of the sound at a given frequency. F_0 denotes the fundamental frequency (approximated by a series of black circles); $F1$ – $F3$ refer to the first three formants; D_f denotes formant dispersion (above, it refers specifically to the dispersion between the second and third formants).

CHAPTER TWO

THE INDEPENDENT AND JOINT EFFECTS OF F_0 AND FN ON LISTENERS' ASSESSMENTS OF BODY-SIZE, MASCULINITY OR FEMININITY, AND ATTRACTIVENESS

2.1 INTRODUCTION

It is well-established that listeners can reliably judge the age and sex of individuals based solely on cues available in the voice (Bruckert *et al.,* 2006; Collins & Missing, 2003; Evans *et al.*, 2006; Childers & Wu, 1991; Owren *et al.*, 2007). Recently, additional attention has been focused on the extent to which listeners can reliably judge other features of speakers from voice cues alone. For example, a growing literature examines listener's ability to assess speaker body-size, masculinity or femininity, and attractiveness.

Some of the rationale for this focus comes from efforts to integrate comparative research on the biological influences of animal communication with humans. In comparative animal work, researchers have been interested in the extent to which vocalizations might convey cues to fitness-related, biological and social dimensions. For many species, assessments of size, masculinity, and attractiveness are essential in evaluations of competitive rivals and potential mates (Darwin, 1871). Larger bodied males, for example, are often the more successful in both the domains of intrasexual competition and intersexual mate-choice, winning more contests and securing more mates, while attractive or highly masculine males or feminine females benefit from higher genetic quality and reproductive success (Alexander *et al.*, 1979; Fitch & Hauser, 2003; Ghazanfar & Rendall, 2008).

As discussed in Chapter One, F_0 and *F*n indirectly correlate with these biosocial dimensions in animals through their connections to anatomical, physiological and hormonal features of the speaker. In most vertebrates, including humans, F_0 is heavily influenced by steroid hormones that affect the size, shape, and tenseness of the vocal folds (Titze, 1989), subsequently affecting their vibratory frequency or pitch (Abitbol *et al.*, 1999; Dabbs & Mallinger, 1999; Evans *et al.*, 2008). Thus, the sizeable, two-fold difference in F_0 between men and women is due to differential exposure to testosterone and estrogen (Jenkins, 1998) and allows listeners to use F_0 as a cue to speaker sex and thus masculinity and femininity. Further, because the same hormones that shape the vocal folds and ultimately determine F_0 may also predict fertility or ovulation in women (Alonso & Rosenfield, 2002; Bryant & Haselton, 2009) and resistance to lowered immunocompetence in men (Folstad and Karter, 1992), F_0 can additionally function to cue speaker quality or attractiveness.

As with F_0, *F*n may also be influenced by and predictive of hormone levels (Bruckert *et al.*, 2006), at least to the extent that hormones influence sexual differentiation and dimorphism in body-size. However, the strongest and most reliable predictor of *F*n is likely the length of an individual’s vocal-tract (Fitch, 1994; Fitch & Hauser, 2003). Thus, we might expect listeners to use both vocal cues in assessments of size, masculinity, and attractiveness with an emphasis on the use of F_0 as a cue to masculinity and attractiveness and *F*n as a cue to size.

Perceptual studies of vocal cueing in humans corroborate these predictions (see Table 1.1). They have been consistent in finding that, when tested independently, lower frequencies of either F_0 (Bruckert *et al.*, 2006; Collins, 2000; Collins & Missing, 2003; Feinberg *et al.*, 2005b, 2006, 2008a; Oguchi & Kikuchi, 1997; Puts, 2005; Riding, Lonsdale & Brown, 2006; Saxton *et al.*, 2006) or Fn (Collins & Missing, 2003; Feinberg *et al.*, 2005b; Puts *et al.*, 2007; van Dommelen & Moxness, 1995; Varosanec-Skaric, 1999) are associated with ratings of larger size and greater masculinity in both sexes and also higher attractiveness in male speakers but lower attractiveness in female speakers. It is well recognized then that both F_0 and Fn affect these social ratings; what remains unknown is precisely how ratings may change across assessments of natural and manipulated voices and across sex of speaker or listener.

In this Chapter, I will further explore these themes in an experiment that attempts to replicate and expand on basic findings concerning the independent influence of F_0 or Fn on listeners' ratings of speaker size, masculinity or femininity, and attractiveness in both naturalistic and manipulated voices. Moreover, I do so using speakers and listeners of both sexes; this is important because females have traditionally been underrepresented in psychoacoustic research. Using both sexes also allows me to account for important differences in the vocal anatomy of men and women and to explore sex-specificity in social attributions.

2.2 METHODS

2.2.1 Voice Recording

The voices used as stimuli in this experiment derived from a speaker database collected two years earlier to minimize the likelihood that participants might recognize the voices of particular speakers. No participant reported recognizing any of the voices used in the experiment. The speaker database included 57 males and 57 females who lacked strong regional accents and whose native language was Canadian English. Speakers were recorded in a sound-controlled room in the Laboratory of Comparative Communication and Cognition at the University of Lethbridge using procedures identical to those used in previous research (for additional details see Rendall *et al.*, 2005, 2007). Briefly, speech samples were collected using an adjustable head-mounted microphone (AKG C420) connected to a preamplifier and computer through a Butterworth antialias filter (Frequency Devices 900/9L8B). The speech material recorded included a list of single-syllable bVt words (e.g., *bit, bet, bat, bait, butt, boot*). Prior to recording, each speaker was given time to familiarize themselves with the list, and to get comfortable saying the words aloud while wearing the headset microphone, while the recordist titrated appropriate signal recording levels. They were subsequently asked to say each item on the list slowly and clearly in a natural speaking voice while a formal recording was made.

2.2.2 Voice Stimulus Selection and Manipulation

From this speaker database, I selected four speakers (2 m, 2 f) whose natural voice pitch and formant values were either relatively low or relatively high for their sex (see Table 2.1 for sample averages and details of acoustic measurements). This provided a natural opportunity for me to test the extent to which inherent variation among speakers representing the extremes of F_0 and Fn for each sex might influence listeners' ratings of size, masculinity and attractiveness.

I selected four additional speakers (2 m, 2 f) with intermediate or mixed values of F_0 and *F*n to create a larger speaker sample. I then performed four independent manipulations of F_0 and *F*n for all eight speakers by either raising or lowering F_0 by 20% while holding *F*n constant; and either raising or lowering *F*n by 10% while holding F_0 constant. The magnitude of these frequency manipulations was chosen so as to parallel those used in previous studies (e.g., Feinberg *et al.*, 2005b; Rendall *et al.*, 2007). In sum, I created a set of five different voice stimuli for each of the eight speakers, each of which was rated on three dimensions (120 trials per participant). Each voice stimulus consisted of five bVt words from the original speech recording (e.g., *bit*, *bet*, *boot*, *bet*, *bat*) with each word separated by 50-ms of silence and standardized to 65 dB (spl).

2.2.3 Participants

Thirty-one females and 30 males completed the experiment. All participants were recruited from the undergraduate community at the University of Lethbridge and received partial course credit. Participants provided informed consent and all but one self-identified as heterosexual. Because notions of masculinity/femininity and attractiveness might vary appreciably between heterosexual and homosexual individuals, the data from the one self-identified homosexual participant were omitted prior to analysis.

2.2.4 Experimental Procedure

Participants completed the experiment privately in a sound-controlled room via a custom computer interface that was designed in ©Runtime Revolution (v. 2.81) to implement the experiment and collect and archive the data. Before testing began, the program presented the participant with a set of instructions outlining the experimental task and procedures.

Participants were instructed that they would hear a series of individual voices played to them one at a time. They heard the voices through Sennheiser HD 280 professional headphones at a comfortable pre-set volume. Their task was to rate each voice on one of three dimensions (either size, masculinity/femininity, or attractiveness) using a 6-point scale. The scale was represented on-screen by a set of six unlabeled radio buttons, three on either side of a mid-point marker that could not be selected. For each dimension, the left and right endpoints of the scale were anchored with text labels that were, respectively, either <*small*> and <*large*>; <*feminine*> and <*masculine*>; or <*unattractive*> and <*attractive*> depending on the dimension being evaluated.

In each trial, one randomly selected voice stimulus was rated on one dimension. The dimension to be rated on a given trial was also randomized within the constraint that each voice stimulus had to be rated three times, once for each of the three dimensions. Each participant completed 120 trials (8 voices x 3 ratings x 5 bVt words) and received three scheduled rest-breaks at equally spaced intervals.

2.2.5 Statistical Analysis

I tested for the effects of F_0 and Fn on participants' ratings of each of the three dimensions using repeated measures Analysis of Variance (rmANOVA) with voice condition as a within-subjects factor and participant sex as a between-subjects factor. Separate analyses were conducted for male and female speakers, for natural and manipulated voices, and for each of three rating dimensions. Post-hoc analyses used the Tukey-Kramer Multiple Comparison Test. Statistical tests were performed using NCSS version 5.1 (Hintze, 1989) using an alpha level of 0.05.

2.3 RESULTS

2.3.1 Natural Voices

As a first pass, I examined listeners' ratings of size, masculinity and attractiveness for the set of four speakers who exemplified naturally low and high extremes in both F_0 and Fn. These tests revealed main effects of voice condition on ratings of all three dimensions (Figure 2.1). Thus, for both male and female raters, speakers of both sexes whose voice had naturally low F_0 and Fn were rated as larger (male speakers (m): $F_{1,59} = 39.56$, $P < 0.05$; female speakers (f): $F_{1,59} = 72.89$, $P < 0.05$) and more masculine (m: $F_{1,59} = 130.19$, $P < 0.05$; f: $F_{1,59} = 267.88$, $P < 0.05$) than speakers with naturally high F_0 and Fn. There were also predictable differences in the average absolute value of these ratings as a function of speaker sex: both male and female listeners rated male speakers as being larger ($M = 4.34$) and more masculine ($M = 4.75$) than female speakers ($M = 2.91$; 2.64).

Ratings of attractiveness also differed as a function of F_0 and Fn, but here there were additional, predictable interactions with the sex of both the speaker and the listener. Thus, speakers with naturally low F_0 and Fn were rated as more attractive if they were male ($F_{1,59} = 15.35$, $P < 0.05$) and less attractive if they were female ($F_{1,59} = 17.76$, $P < 0.05$), and these effects held differentially for male and female listeners. Hence, only female listeners rated male speakers with low F_0 and Fn as more attractive than those with high F_0 and Fn ($F_{1,59} = 14.08$, $P < 0.05$), while only male listeners rated female speakers with low F_0 and Fn as less attractive than those with high F_0 and Fn ($F_{1,59} = 7.11$, $P < 0.05$). In other words, for both male and female listeners, ratings of the attractiveness of same-sex speakers did not vary as a function of voice F_0 and Fn, perhaps because rating the attractiveness of same-sex speakers is a more difficult or unnatural task.

2.3.2 Manipulated Voices

The outcomes for the joint effects of either low or high F_0 and Fn in natural voices were corroborated in analyses of listeners' ratings of manipulated voices where F_0 or Fn were independently and systematically increased or decreased (Figure 2.2). Male and female speakers whose voices had been lowered in F_0 were rated by both sexes as larger (m: $F_{1, 59} = 53.48$, $P < 0.05$; f: $F_{1, 59} = 18.86$, $P < 0.05$) and more masculine (m: $F_{1, 59} = 233.8$, $P < 0.05$; f: $F_{1, 59} = 246.38$, $P < 0.05$) compared to speakers whose F_0 had been increased. Similarly, male and female speakers whose voices had been lowered in Fn were rated as larger (m: $F_{1, 59} = 158.92$, $P < 0.05$; f: $F_{1, 59} = 193.75$, $P < 0.05$) and more masculine (m: $F_{1, 59} = 180.43$, $P < 0.05$; f: $F_{1, 59} = 120.14$, $P < 0.05$) than speakers with raised Fn, and this was true for both male and female listeners. Once again, both male and female listeners rated male speakers as being, on average, larger ($M = 4.32$) and more masculine ($M = 4.61$) than female speakers ($M = 2.87$; 2.56).

Ratings of attractiveness also differed as a function of my independent manipulations of F_0 or Fn and in ways that again interacted with the sex of both speaker and listener. Thus, speakers whose F_0 or Fn had been lowered, compared to those whose F_0 or Fn had been raised, were rated as more attractive if they were male (F_0: $F_{1, 59} = 10.12$, $P < 0.05$; Fn: $F_{1, 59} = 6.62$, $P < 0.05$) and less attractive if they were female (F_0: $F_{1, 59} = 45.21$, $P < 0.05$; Fn: $F_{1, 59} = 47.64$, $P < 0.05$). As before, these effects of F_0 and Fn on attractiveness ratings held only for female (and not male) listeners rating male voices ($F_{1, 59} = 14.75$, $P < 0.05$) but for both male and female listeners rating female voices.

2.4 DISCUSSION

Listeners consistently judged speakers with either low F_0 or low Fn, or both, as larger and more masculine, and either attractive if the speaker was male or unattractive if female. These outcomes confirm previous results showing reliable differentiation of speaker sex based on pitch and formant profiles (Childers & Wu, 1991; Coleman, 1976**;** Jenkins, 1998; Owren *et al.,* 2007) as well as a growing body of evidence implicating the same two voice features in assessments of attractiveness (Bruckert *et al.,* 2006; Collins, 2000; Collins & Missing, 2003; Feinberg *et al.,* 2005a, 2005b, 2006, 2008a; Hughes *et al.,* 2002; Hughes and Gallup, 2008; Puts, 2005; Saxton *et al.,* 2006). Previous studies have also found that male speakers with low F_0 or low Fn are consistently rated larger and more masculine (Collins, 2000; Evans *et al.,* 2006; Feinberg *et al.,* 2005b; González, 2006). My results further indicate that the same low F_0 or low Fn features that elicit ratings of larger size and masculinity in men occur in female speakers as well. The latter outcome suggests that the perceptual algorithm listeners use to assess body-size and masculinity from the voice may be a general one that is not differentiated by the sex of the speaker.

In apparent contradiction to my own findings, two previous studies have failed to find an effect of Fn on females' assessments of the attractiveness (Feinberg *et al.,* 2005b) or pleasantness (Bruckert *et al.,* 2006) of male speakers, but possibly for very sensible reasons. Feinberg *et al's* study used only a 5% shift in formant frequencies relative to baseline and so it is possible that many of the formant contrasts fell below listeners' threshold discrimination abilities (for thresholds see Chapter 3, Part I). Bruckert *et al's* study involved assessments of 'pleasantness' as opposed to 'attractiveness' *per se* and the two constructs might be meaningfully different. At the same time, the study used naturalistic speech sequences in which the more protracted F_0 intonation patterns across the word sequences of different speakers were allowed to vary (i.e., involve variably rising or falling pitch contours across words). Ultimately, these variable F_0 intonation patterns were found to influence assessments of vocal pleasantness more than did the absolute frequencies of either F_0 or Fn.

Furthermore, several other studies have found that low F_0 voices, in addition to low Fn voices, are perceived as belonging to larger bodied persons (Collins, 2000; Feinberg *et al.*, 2005b; van Dommelen & Moxness, 1995), despite there being no real correlation between F_0 and height (see Table 1.2). There are many possible explanations for this outcome. For instance, the misattribution of low F_0 to large size may simply reflect a generalization from across sexes (where there are marked differences in the F_0 of males and females) to within sexes or perhaps from inanimate objects in the real world to people, where larger objects tend to produce deeper pitched sounds. It may also be the case that listeners use F_0 to gauge size because of its correlations with hormone levels and masculinity, a relationship they then extrapolate (erroneously but intuitively) to body-size. Alternatively, as proposed by Fitch (1999, 2000), men may actively lower their vocal pitch so as to falsely advertise a larger body-size; this hypothesis may also explain the huge disparity between male and female average F_0. Similarly, F_0 may simply be more audible or salient than Fn, and thus, more readily abstracted to assessments of body-size. A fifth interpretation, proposed by González & Oliver (2004), appeals to the idea that body-size information is not located in Fn alone, but rather dispersed across several speech parameters, including F_0. Lastly, there is some evidence to suggest that although F_0 does not correlate with height, it may correlate with weight (Collins & Missing, 2003; Evans *et al.,* 2006) or the shape of one's body (e.g., shoulder-hip ratio, Evans *et al.,* 2006; Hughes *et al.,* 2004), suggesting that there may be some logic behind listeners' extrapolations of F_0 to assessments of body-size.

Although this study reports consistent effects of F_0 or Fn on listeners' assessments of size, masculinity and attractiveness, these patterns do not yet allow us to say which of F_0 or Fn might be most responsible for the effects. According to conventional theories of speech production (Fant 1960; Müller, 1858), although voice F_0 and Fn can sometimes be coupled, they are often at least partially independent of one another. Empirical studies, and everyday experience, confirm this independence such that there can be considerable independent variation in F_0 and Fn both within and across utterances by the same individual and in the baseline average values of these features in different individuals. For example, there can be individuals with relatively low F_0 but high Fn (and vice versa) for their sex as well as individuals exemplifying many possible combinations of average F_0 and Fn values between the extremes. As a result, it cannot be assumed that, for any given speaker, voice F_0 and Fn will necessarily be highly correlated and thus provide wholly redundant cues to size, masculinity and attractiveness. The corollary is that listeners might differentially weight voice F_0 and Fn in making voice-based social assessments of others. I examine this possibility in the next Chapter.

Table 2.1. Values of F_0 and Fn for male and female speakers. Values shown reflect acoustic measurements performed using ©PRAAT software (version 4.6, Boersma, 2001). Measurements of F_0 and Fn were taken from the central, steady-state portions of the vowel sound in each of five bVt words and then averaged within speakers.

Baseline Frequency (Hz)	Male Speakers[a]				
	1	2	3	4	
	Low F_0Fn	High F_0Fn	Mixed F_0Fn	Mixed F_0Fn	Mean
F_0	93	138	82	134	**113**
$F1$	412	410	399	472	**468**
$F2$	1471	1922	1892	1697	**1431**
$F3$	2467	2772	2597	2562	**2505**
$F4$	3433	3814	3494	3560	**3468**

Baseline Frequency (Hz)	Female Speakers[a]				
	1	2	3	4	
	Low F_0Fn	High F_0Fn	Mixed F_0Fn	Mixed F_0Fn	Mean
F_0	132	234	165	178	**204**
$F1$	452	535	483	526	**583**
$F2$	1633	2122	2055	2051	**1747**
$F3$	2612	3054	3067	3034	**2915**
$F4$	3587	4157	4224	3872	**4089**

a. Values in columns 1-4 are for speakers used in the current experiment, while those in the column labeled ‘mean’ are included for comparison and represent means from a sample of 34 male and 34 female speakers as reported in Rendall *et al.*, 2005.

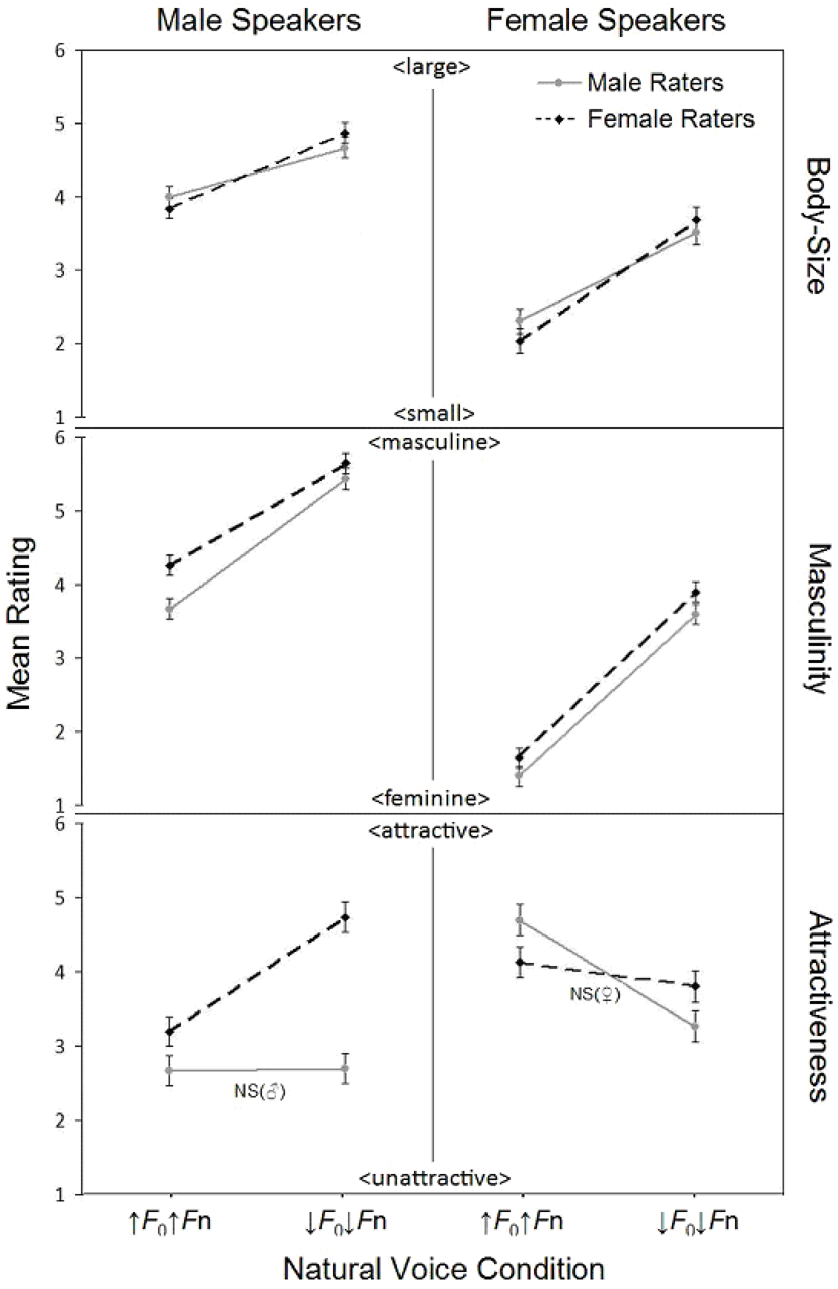


Figure 2.1. Mean (±SE) ratings of body-size, masculinity (femininity), and attractiveness for male and female speakers with naturally high F_0 and Fn compared to those with naturally low F_0 and Fn (↑F_0↑Fn vs.↓F_0↓Fn). NS: $P > 0.05$.

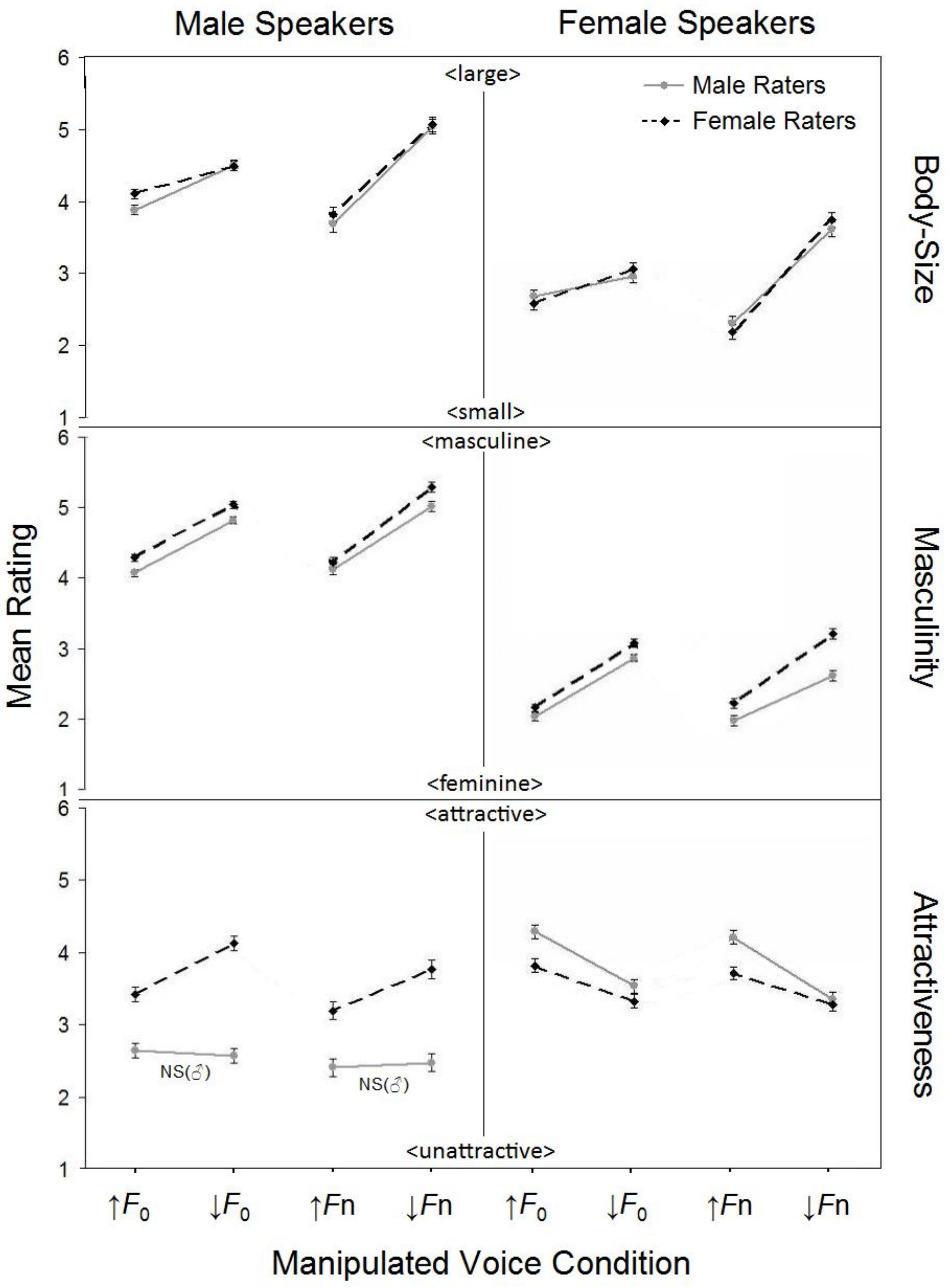


Figure 2.2. Mean (±SE) ratings of body-size, masculinity (femininity), and attractiveness for male and female speakers with manipulated high versus low F_0 (↑F_0 vs.↓F_0) or high versus low *F*n (↑*F*n vs.↓*F*n). NS: $P > 0.05$.

CHAPTER THREE

RELATIVE TRACKING OF FN VERSUS F_0 IN LISTENERS' VOICE-BASED BIOSOCIAL ASSESSMENTS

3.1 INTRODUCTION

Low frequencies of voice fundamental (F_0) or formants (Fn) reliably elicit from listeners attributions of large size and masculinity, whether they occur in male or female voices (Collins, 2000; Feinberg *et al.*, 2005b). Low voice frequencies are also attractive when they occur in male voices (Collins, 2000; Saxton *et al.*, 2006), while the opposite is true of female voices where higher frequencies are more attractive (Collins & Missing, 2003). It might be that deeper voiced individuals tend to be taller and have higher levels of testosterone, and hence that men with these characteristics tend also to be more attractive; however, size and masculinity are not equally correlated with F_0 and Fn. Thus, while Fn correlates with height, F_0 does not (Rendall *et al.*, 2005); and while F_0 correlates with testosterone levels (Dabbs & Mallinger, 1999), Fn is less likely to (but see Bruckert *et al.*, 2006).

As discussed in earlier Chapters, the argument that F_0 and Fn cue different speaker dimensions follows a proposal made by Fitch (1994, 2000; Fitch & Hauser, 2003) with respect to voice-based assessments of body-size. Fitch emphasized that while relaxed voice pitch is determined largely by the length and mass of the vocal folds, formant frequencies are determined by the length of the vocal-tract that, in turn, is constrained by the length and size of surrounding bony structures (e.g., neck, pharynx, oro-pharynx and oral cavity). As a result, the length of the vocal-tract is anatomically more constrained than is the length (and mass) of the vocal folds, with the former being more obligatorily connected to overall body-size. Hence, Fn is the more honest (*sensu* Zahavi, 1975) and reliable cue to size (*qua* height), and should, accordingly, be tracked over F_0 in assessments of size. Likewise, F_0's correlations with testosterone and estrogen suggests that it should be tracked over Fn is assessments of masculinity or femininity. Yet, listeners' assessments of size and masculinity have been shown to rely on cues from both vocal features, rather than only the more reliable one.

It may be the case that cues from F_0 and Fn are virtually redundant, where one or the other functions as a backup system of cues to speaker quality. However, because F_0 and Fn are relatively independent features of the voice (Fant 1960; Müller, 1848) and, at the same time, differentially correlated with certain aspects of human anatomy and physiology, it is more likely that F_0 and Fn actually cue different dimensions of a speaker. Consequently, although it may be true that listeners rely on cues from both F_0 and Fn to assess both size and masculinity, as well as attractiveness, the *degree* to which they rely on each vocal feature may differ across different assessments based on predictable variation in the degree to which each vocal feature correlates with each dimension. In other words, although listeners may rely on both F_0 and Fn to assess size, they may give more precedence to cues from Fn; and similarly for assessments of masculinity, listeners may rely on both F_0 and Fn but give more precedence to cues from F_0. This prediction cannot be tested by studying the effects of either vocal feature in isolation. Instead, a more detailed analysis of the *relative* effects of these vocal features on listeners' assessments is necessary.

Thus, this Chapter attempts to titrate the relative salience of F_0 and Fn and to establish which of these vocal features most heavily influences listeners' voice-based judgments of size as well as masculinity and attractiveness. Following Fitch, I predict that assessments of body-size ought to be more heavily influenced by Fn than F_0. In contrast, because F_0 is arguably more closely related to testosterone and estrogen than Fn and is also more sexually dimorphic, I predict that assessments of masculinity ought to be more heavily influenced by F_0 than Fn. It is more difficult to speculate which of F_0 or Fn listeners ought to track in assessments of attractiveness, as both body-size (cues from Fn) and masculinity/femininity (cues from F_0) have been shown to affect intersexual ratings of attractiveness (Yates, Edman & Aruguete, 2004; Feinberg, DeBruine, Jones & Little, 2008b). Nevertheless, supposing that attractiveness is closely related to speaker sex and, given that speaker sex is better cued by F_0 than Fn, I predict that listeners will more readily track F_0 over Fn in assessments of attractiveness.

PART I. FREQUENCY DISCRIMINATION THRESHOLDS FOR F_0 AND FN

In order to test the relative effects of F_0 versus Fn, an important preliminary step was to establish discrimination thresholds for frequency differences in each of the vocal features independently. Although there is a substantial literature on frequency discrimination thresholds (or just-noticeable differences, JNDs) in humans, much of it focuses on the discrimination of either pure-tone stimuli (Klinge & Klump, 2009; Wier, Jesteadt, & Green, 1977), or broad-band stimuli with emphasized center-frequencies to simulate formants, rather than naturalistic speech material *per se* (Kewley-Port & Watson, 1994; Mermelstein, 1978). There is comparatively little research on discrimination thresholds of either F_0 or Fn in naturalistic speech (cf. Puts *et al.*, 2007). Because our interest is ultimately to determine how voice-based social judgments are variably affected by pitch and formants as they are embedded in naturalistic speech, I first sought to establish frequency discrimination thresholds for these two features and for the particular voice samples I was using.

3.2 METHODS

3.2.1 Voice Stimuli and Analysis

The voices used as stimuli in this experiment derived from the same speaker database and involved the same word material (single-syllable words in bVt context) as described for the previous experiment (Chapter 2). From this database I selected eight new speakers (4 m, 4 f) and for each one used an identical set of four words. To facilitate construction of experimental stimuli, I first obtained measurements of voice F_0 and Fn for each speaker. Measurements were performed using ©PRAAT software (v. 4.6, Boersma, 2001) and were taken from the central, steady-state portion of the vowel in each of the four words. Pitch measurements were obtained using PRAAT's pitch-tracking function with a range setting of 75 Hz to 300 Hz. Formant measurements were obtained using PRAAT's formant-tracking function and interactive modification of the number of formants identified (i.e., 4±2) depending on speaker sex and visual evaluation of the fit of putative formant tracks to manifest spectral content of the vowels. The resulting values for F_0 and Fn for different vowels were then averaged within speakers to establish baseline F_0 and Fn values for each speaker. The mean baseline F_0 for our sample of male speakers was 120 Hz (range: 114-127 Hz) with mean values for F1 - F4 of 353, 1635, 2690, and 3522 Hz, respectively. The mean baseline F_0 for our sample of female speakers was 207 Hz (range:

172- 242 Hz) with mean formant values of 443, 1914, 3006, and 4068 Hz. These values agree well with those of previous samples of American English speakers (e.g., Stevens, 1998; Bachorowski and Owren 1999; Rendall *et al*., 2005).

3.2.2 Stimulus Construction

Experimental stimuli consisted of a pairing of two sets of the same four bVt words (*boat, beat, book, bait*) spoken by the same individual with a 600-ms silent interval between each word set and a 50-ms silent interval between each word within a set. In test-trials, each stimulus contained the original, unmanipulated recording (baseline condition) of the four-word set by a given speaker followed by a repetition of the same word set by the same speaker but with either F_0 or Fn increased by 1-10% relative to that speaker's mean baseline values (manipulated conditions). In catch-trials, each stimulus contained two presentations of the same word set by a given speaker, either two presentations of the baseline condition or two presentations of one of the manipulated conditions (1-10%). Frequency manipulations were performed using PRAAT either by specifying a new absolute pitch median for each word (in the case of pitch modifications) or a proportional shift-ratio to apply to all formants (in the case of format modifications). All experimental stimuli were standardized for length (at 5 s) and amplitude (65 dB).

3.2.3 Participants

Thirty-seven females and 25 males completed the experiment. All participants were recruited from the University of Lethbridge undergraduate community and received partial course credit.

3.2.4. Experimental Procedure

Before testing began, participants were assigned randomly to one of four testing groups that involved making frequency discriminations in either pitch or in formants and in the voices of one or the other of two sets of four speakers (set A or set B, each containing 2 m and 2 f speakers). Otherwise, the general procedures for this experiment were the same as those for the previous one. Participants once again completed the experiment in a sound-controlled room via a custom computer interface. The program first presented the participant with a set of instructions outlining the frequency discrimination task. Participants were instructed that they would hear a series of voice-comparisons that would involve two presentations of the same set of four words spoken by the same person. They were informed that the two repetitions might be identical to one another in frequency or slightly different; their task was simply to indicate whether the repetitions were the exact <*same*> or whether they were <*different*> by clicking on the appropriate button on the computer screen bearing that label. Participants were invited to watch a demonstration-trial before beginning the formal experiment. There were no time restrictions and they were free to replay any given stimulus one time using a <*replay*> button on the computer screen. Each participant received 224 trials subdivided into four blocks. Each block contained ten test trails and four catch trials for each of four speakers (n = 56 trials/block), with the order of these trials randomized.

3.2.5 Statistical Analysis

I used rmANOVA tests to examine variation in listeners' discrimination performance as a function of subject sex, speaker set (A or B), and frequency-difference condition (1-10%). Separate tests were run for discriminations of F_0 and Fn and for male and female speakers. All ANOVAs were a mixed-design with listener sex and speaker set included as between-subjects factors and frequency-difference condition as a within-subjects factor. Statistical tests used an alpha level of 0.05.

I also used a signal detection paradigm to calculate d-prime (d') scores for each incremental frequency change in F_0 or Fn from 1 to 10%. D-prime is a standard metric used to assess listener sensitivity to stimulus variation while controlling for individual response biases (Macmillan and Creelman, 2005). It can be estimated as the standardized difference between the means of the distributions of listeners' correct responses on test-trials (hits) versus their incorrect responses on catch-trials (false alarms). I used a d' value of 1 as a criterion for recognizing a reliable degree of discrimination performance, or a just-noticeable difference (JND).

3.3 RESULTS

Overall, there were significant main effects of the frequency-difference conditions on the proportion of correct discriminations made by listeners for both F_0 and Fn in both male and female speakers (male (m) F_0: $F_{9, 243} = 38.25$, $P < 0.05$; female (f) F_0: $F_{9, 243} = 42.67$, $P < 0.05$; m Fn: $F_{9, 243} = 65.80$, $P < 0.05$; f Fn: $F_{9, 243} = 70.19$, $P < 0.05$). There were no effects of listener sex on discrimination performance (m for F_0: $F_{1, 27} = 1.75$ $P = 0.2$; f for F_0: $F_{1, 27} = 0.02$ $P = 0.88$; m for Fn: $F_{1, 27} = 0.08$ $P = 0.79$; f for Fn: $F_{1, 27} = 0.79$ $P = 0.38$). Thus, men and women performed equally well when discriminating F_0 and Fn in voices of the same or of the opposite sex.

In F_0 discriminations for male voices, there was also a significant main effect of speaker set ($F_{1,\,27} = 10.02$, $P < 0.05$) and an interaction between speaker set and frequency dimension ($F_{9,\,243} = 4.25$, $P < 0.05$) where differences in F_0 were more easily discriminated in male speakers in set B than in set A. This outcome suggests that the discriminability of voice features may vary somewhat from speaker-to-speaker (see Table 3.1). There were no other significant effects.

Predictably, listener discrimination of frequency-differences in F_0 and *F*n improved steadily as the magnitude of the frequency difference between the baseline and test voice stimuli increased from one to ten percent. This outcome is shown in Table 3.1 where listeners' discrimination performance for each speaker and for all frequency-difference conditions are summarized as *d'* scores. Inspection of these scores confirms variation in the discriminability of F_0 and *F*n differences from speaker-to-speaker. However, it also shows that *d'* scores generally increased steadily as the magnitude of the frequency difference increased. Moreover, there was some consistency in the point at which *d'* scores exceeded and remained above the criterion value of 1. For both F_0 and *F*n, that point was generally at frequency-differences of 4% to 7%.

Figure 3.1 displays these patterns graphically, where *d'* scores for each frequency-difference condition are averaged across listeners (within speaker sex) and plotted separately for male and female speakers discriminating either F_o or *F*n. The pattern of steadily increasing discrimination performance is made clearer as is the consistency in the point at which this exceeds and remains above criterion. From these averages, I was able to delineate reliable discrimination performance (1 JND) for pitch differences in both male and female voices at 6%, and for formant differences at 6% for male voices and at 5% for female voices.

3.4 DISCUSSION

As already noted, our discrimination task using naturalistic speech material is not entirely comparable to many past studies. Hence, it is difficult to compare outcomes directly. For example, previous studies using pure tones or tone complexes (Moore, 1973; Klinge & Klump, 2009; Sinnott, Brown, & Brown, 1992; Sinnott, Owren and Peterson, 1987; Wier, Jesteadt, & Green, 1977) have generally reported JNDs lower (1-2%) than those I found for the discrimination of F_0 in naturalistic speech. This difference might prove robust and point to higher discrimination thresholds generally for F_0 variation in naturalistic speech compared to synthetic tones or tone complexes. Our stimuli also involved a sequence of multiple words, rather than only a single sound, and thus included some variation among words in F_0 (as well as in formants). This additional variation might also have contributed to higher discrimination thresholds in our task compared to the discrimination of pure tones. In this case, the stimuli used for discrimination were sufficiently different in our experiment compared to most past studies, that it is difficult to fully reconcile the differences.

Other comparisons are more appropriate and promising. Thus, previous studies testing formant discrimination thresholds using synthesized vowel sounds have reported JNDs more comparable to those that I found, in the range of 1.5 - 9 % (reviewed in Kewley-Port, Li, Zheng, & Neel, 1996) with variation among studies attributed to many subtle methodological differences between them including: whether or not subjects were trained prior to testing (Kewley-Port & Watson, 1994; Kewley-Port, 1995); whether vowels were presented in isolation or consonantal context (Ives, Smith & Patterson, 2005; Kewley-Port & Watson, 1994; Mermelstein, 1978); and whether only one or many different speakers were used and whether they were male or female and involved either natural or synthesized voices (Kewley-Port, 1995; Kewley-Port & Watson, 1994; Smith *et al.*, 2005). Although none of these studies involved short stretches of naturalistic speech or such a large sample of both speakers and listeners, the results are relatively consistent and encompass the discrimination thresholds of 5-6% that I observed.

The most comparable previous study to mine was by Puts, Hodges, Cardenas & Gaulin (2007) who reported comparable JNDs in F_0 and formants of approximately 7% and 4%, respectively, for full-sentence stimuli from male speakers.

PART II. THE RELATIVE TRACKING OF F_0 VERSUS *F*n IN VOICE-BASED BIOSOCIAL JUDGEMENTS

Having established JNDs for discrimination of F_0 and *F*n for my speaker sample, I conducted an experiment to test which of F_0 or *F*n listeners might weigh more heavily when making voice-based social judgments. Put simply, I created experimental voice stimuli that would mimic natural speakers whose F_0 and *F*n features provided conflicting cues to size, masculinity and attractiveness because they combine relatively low F_0 with relatively high *F*n (or vice versa). To conduct this experiment properly, the discordance between the two acoustic features was controlled by using perceptually equivalent manipulations in F_0 and Fn (JNDs).

If listeners attend to and weight F_0 and Fn cues equally in assessing size, masculinity and attractiveness, then I predict that there should be no consistent variation in their ratings of such stimuli because the conflicting effects of the F_0 and Fn cues would cancel each other out. In contrast, if listeners attend to and weigh one voice feature more than the other, then there should be consistent variation in their ratings with the direction of the effects indicating which of the two features was the more salient. As a possible example, because large size is consistently associated with lower-frequency voices, if voice stimuli combining high F_0 with low Fn ($\uparrow F_0 \downarrow F$n) were consistently rated larger by comparison to those combining low F_0 and high Fn ($\downarrow F_0 \uparrow F$n), then Fn cues would be being weighed more heavily than F_0 cues.

3.5 METHODS

3.5.1. Voice Stimuli and Analysis

The voices used as stimuli in this experiment once again derived from the same speaker database and involved the same word material as described for the previous two experiments. From this database, I selected 20 new speakers (10 m, 10 f) and for each speaker an identical set of five words. Once again, I first obtained measurements of voice F_0 and Fn for each speaker using PRAAT pitch-tracking and formant-tracking functions. The resulting values for the vowels in each of the five words were averaged within speakers to establish baseline F_0 and Fn values for each speaker. For male speakers in this sample, the mean F_0 was 98 Hz (range: 82-117 Hz) and the mean values for $F1$ – $F4$

were 469, 1619, 2584, and 3511 Hz, respectively. For female speakers, the mean F_0 was 204 Hz (range: 178-236 Hz) and the mean formant values were 586, 1949, 2969, and 4047 Hz, respectively.

3.5.2. Stimulus Construction

Experimental stimuli for this study involved the same sample of single-syllable words used previously, this time specifically the words *bet, butt, bite, beat, book* with each word separated by a 50-ms silent interval and all standardized to 65 dB. I used the natural (unmanipulated) sample of these words for each speaker as a baseline condition and developed four additional experimental conditions that involved modifications of F_0 and *F*n. These modifications entailed putting pitch and formant cues in conflict with each other by raising F_0 while lowering *F*n or by lowering F_0 while raising *F*n. Modifications to F_0 and *F*n were performed in increments designed to be equally perceptually discriminable to listeners and hence were done in JND increments guided by the results of the previous experiment. Those results indicated that, for both male and female voices, the JND for F_0 was 6%, while the JND for *F*n was also 6% for male voices and 5% for female voices. In previous studies of F_0 and *F*n effects, subtle frequency manipulations in the range of one JND to one or other vocal feature have failed to produce consistent results (Feinberg *et al.*, 2005b) compared to shifts of two or more JNDs (Puts, 2005) and our own pilot testing confirmed this pattern. Hence, I chose to create experimental stimuli using F_0 and *F*n shifts of two and three JNDs. Such shifts would be more perceivable to listeners but still result in parameters for both voice features within the normal range for each sex. For example, modifications to baseline F_0 values for my male speakers by two and three JNDs yielded F_0 values of 67 – 138 Hz; while modifications of two and three JNDs to baseline F_0 values for my female speakers yielded F_0 values of 145-279 Hz.

The four experimental conditions thus entailed stimuli in which either the pitch was raised and the formants were lowered by two or three JNDs (i.e., $\uparrow 2F_0 \downarrow 2F\text{n}$; $\uparrow 3F_0 \downarrow 3F\text{n}$), or the pitch was lowered and the formants were raised by two or three JNDs (i.e., $\downarrow 2F_0 \uparrow 2F\text{n}$; $\downarrow 3F_0 \uparrow 3F\text{n}$). Manipulations were performed using PRAAT by delineating a new absolute pitch median for each word and applying a proportional shift-ratio to all formants.

3.5.3. Participants

Thirty-two females and thirty-six males completed this experiment. All participants were recruited from the University of Lethbridge undergraduate community and received course credit. This sample includes only those participants who self-reported as heterosexual.

3.5.4. Experimental Procedure

Before testing began, participants for this experiment were assigned randomly to one of two versions of the experiment. In Version A (*n* = 18 m, 16 f), voice stimuli and rating dimensions were completely randomized, within the constraint that each voice stimulus had to be rated three times, once for each of the three dimensions. In Version B (*n* = 18 m, 16 f), rating dimensions were blocked, with each block of trials involving the same dimension (either size, masculinity/femininity or attractiveness). The order of these blocks was randomized across participants. The order of voice stimuli presented within each block was also randomized within and between participants. In this experiment, listeners only rated speakers of the opposite sex. Otherwise, both versions of the experiment included the complete speaker set (*n* = 10 speakers of one or the other sex), the same set of baseline and experimental conditions for each speaker (*n* = 5 conditions per speaker) and the same three rating dimensions, for a total of 150 trials per subject.

All other procedures for this experiment were identical to those used in the first experiment (Chapter 2). On each trial, participants were presented with a single voice stimulus of the opposite-sex with the opportunity to replay the stimulus once. Their task was to rate each voice on one of the three dimensions using the same 6-pt scale with no midpoint described in Chapter Two. Subjects were given an unlimited amount of time to complete the experiment. The only difference in the task in this experiment was that participants' response times (RT's) were recorded. Rest breaks were scheduled at 50-trial intervals.

3.5.5. Statistical Analysis

Statistical testing paralleled that for the first experiment except that experimental version was included as a between-subjects factor and both rating dimension and experimental frequency condition were within-subjects factors in rmAVNOVA tests. Trials on which participant response time was less than 50-ms were removed prior to analysis.

3.6 RESULTS

For female listeners responding to male speakers, there were significant main effects of experimental frequency condition ($F_{4,\,120} = 160.6$, $P < 0.05$) and rating dimension ($F_{2,\,60}$

$= 90.17$, $P < 0.05$) as well as a significant interaction between the two ($F_{8,\,240} = 85.76$, $P < 0.05$). Post-hoc analyses showed that females' ratings of all three dimensions were significantly greater for the voice conditions where Fn was lowered while F_0 was raised either by 2 JNDs or by 3 JNDs compared to the corresponding voice conditions where F_0 was lowered and Fn was raised (size: $\uparrow 2F_0 \downarrow 2F\mathrm{n}$ $M = 4.88$; $\downarrow 2F_0 \uparrow 2F\mathrm{n}$ $M = 3.05$; $\uparrow 3F_0 \downarrow 3F\mathrm{n}$ $M = 5.53$; $\downarrow 3F_0 \uparrow 3F\mathrm{n}$ $M = 2.44$; masculinity: $\uparrow 2F_0 \downarrow 2F\mathrm{n}$ $M = 5$; $\downarrow 2F_0 \uparrow 2F\mathrm{n}$ $M = 3.44$; $\uparrow 3F_0 \downarrow 3F\mathrm{n}$ $M = 5.47$; $\downarrow 3F_0 \uparrow 3F\mathrm{n}$ $M = 3.02$; attractiveness: $\uparrow 2F_0 \downarrow 2F\mathrm{n}$ $M = 3.58$; $\downarrow 2F_0 \uparrow 2F\mathrm{n}$ $M = 2.93$; $\uparrow 3F_0 \downarrow 3F\mathrm{n}$ $M = 2.53$; $\downarrow 3F_0 \uparrow 3F\mathrm{n}$ $M = 1.86$). There was no effect of Version on females' ratings ($F_{1,\,30} = 2.18$, $P = 0.15$), implying that blocking trials by rating dimension did not affect listeners' perceptions of the various dimensions.

Similarly, for male listeners responding to female speakers, there were significant effects of experimental frequency condition ($F_{4,\,136} = 46.72$, $P < 0.05$), rating dimension ($F_{2,\,68} = 6.72$, $P < 0.05$), and a significant interaction between the two ($F_{8,\,272} = 112.72$, $P < 0.05$). Likewise, post-hoc analyses showed that males' ratings of size and masculinity were significantly greater for the voice conditions where Fn was lowered while F_0 was raised either by 2 JNDs or by 3 JNDs compared to the corresponding voice conditions where F_0 was lowered and Fn was raised (size: $\uparrow 2F_0 \downarrow 2F\mathrm{n}$ $M = 4.05$; $\downarrow 2F_0 \uparrow 2F\mathrm{n}$ $M = 2.68$; $\uparrow 3F_0 \downarrow 3F\mathrm{n}$ $M = 4.86$; $\downarrow 3F_0 \uparrow 3F\mathrm{n}$ $M = 2.47$; masculinity: $\uparrow 2F_0 \downarrow 2F\mathrm{n}$ $M= 3.34$; $\downarrow 2F_0 \uparrow 2F\mathrm{n}$ $M = 2.7$; $\uparrow 3F_0 \downarrow 3F\mathrm{n}$ $M = 4.14$; $\downarrow 3F_0 \uparrow 3F\mathrm{n}$ $M = 2.88$). Males' ratings of female attractiveness were the reverse of this pattern where their ratings were significantly greater for the voice conditions where Fn was raised while F_0 was lowered either by 2 JNDs or 3 JNDs compared to the corresponding voice conditions where F_0 was raised while Fn was lowered (attractiveness: $\downarrow 2F_0 \uparrow 2F\mathrm{n}$ $M = 3.85$; $\uparrow 2F_0 \downarrow 2F\mathrm{n}$ $M = 3.15$; $\downarrow 3F_0 \uparrow 3F\mathrm{n}$ $M = 3.1$;

↑3F_0↓3Fn $M = 2.05$). There was no effect of Version on males' ratings ($F_{1, 34} = 0.74$, $P = 0.4$).

3.7 DISCUSSION

Listeners of both sexes consistently rated speakers of the opposite sex as larger and more masculine (or less feminine) in experimental conditions in which Fn had been lowered (while F_0 had been raised). Speakers with lowered Fn (but raised F_0) were also rated more attractive if they were male speakers, and less attractive if they were female speakers.

While these results indicate that listeners' ratings of size, masculinity and attractiveness were affected more by Fn, previous studies have reported stronger effects for F_0 in ratings of speaker size (Collins 2000; González & Oliver, 2004), sex (Coleman, 1976), masculinity (Collins, 2000), and speaker attractiveness (Collins, 2000). The primacy of F_0 over Fn reported in these studies, however, reflects only its greater predictability of listener's ratings of *natural* voices, where unmanipulated vocal F_0 and Fn maintain their perceptual disparity. Unlike my own, then, this body of work does not address which vocal feature is most prominent after having controlled for their variable perceptual salience or when placed in conflict with one another; it is nevertheless consistent in pointing to the pre-eminence of voice F_0 in such social ratings of *natural* voices and that outcome is certainly intuitive. After all, the voice differences typically associated with these distinctions in the natural classes of speakers that exemplify them (i.e., males and females) are much larger for F_0 than for Fn: there is a nearly two-fold difference in voice F_0 between adult males and females, while the Fn differences between the sexes are on the order of only 10-20% (Hillenbrand, Getty, Clark, Wheeler, 1995; Wu & Childers, 1991; see Table 2.1). Hence, using F_0 more than Fn cues to judge speaker sex, and the closely associated dimensions of masculinity/femininity and possibly also body-size, might be simply taking advantage of the much greater natural variation, and thus discriminability, of F_0 compared to Fn that mark these distinctions between males and females.

At the same time, this point might also highlight the value of controlling the magnitude and relative discriminability of F_0 and Fn differences used in experimental tests of their relative salience. By first establishing equally discriminable differences in F_0 and Fn for my speaker sample and then putting the two cues in conflict by equally discriminable amounts, I was arguably able to more definitively test their relative effects. And, at least under these conditions, my results suggest that listeners might weigh Fn cues more than F_0 cues in biosocial assessments.

Ultimately, this outcome does not actually contradict the results of previous studies so much as complement them in showing that, under controlled conditions in which differences in the two voice features are made equally perceptually discriminable, the typical priority of F_0 compared to Fn cues obtained for many naturalistic speaker contrasts can be reversed. Notably, one other recent study established and used JNDs for F_0 and Fn in the synthesis of experimental stimuli and found that, while both low F_0 and low Fn elicited high ratings of social and physical dominance in male speakers, the effects were greater for Fn than for F_0 (Puts *et al.*, 2007).

If true and generalizable in future work, the apparent priority of Fn over F_0 in contexts like those tested here (and in Puts *et al.*, 2007) might arise for one of four reasons. First, it might be that the Fn variation in our experiment was simply more discriminable. We explicitly strived to avoid this possibility by the precautions of first establishing discrimination thresholds for both features and then using stimuli that put the two features in conflict by equivalent JND increments. Nevertheless, while equivalent in relative terms, a 1 JND shift (6%) in Fn obviously involves a much larger absolute frequency shift than does a 1 JND shift (6%) in F_0 because the baseline frequency from which such shifts are made is much higher for all formants than it is for F_0. Of course, it is exactly this absolute-relative frequency distinction that my JND precautions were meant to control and, so, in emphasizing this distinction again, I am not now trying to undermine the logic of my own experiment or somehow question the broader logic behind discrimination threshold testing generally and the establishment of JNDs. However, I am raising the possibility that thresholds for discriminating variation in Fn and F_0 might change consistently in the context of complex stimuli containing variation in both features compared to those established for variation in each feature in isolation. Indeed, Kewley-Port and Watson (1994) reported that discrimination thresholds for formant frequencies did vary slightly as a function of variation in the F_0 of the stimuli in which they were jointly embedded. It is not clear exactly if or how a shifting discrimination threshold effect like this might *consistently* favour Fn over F_0 cues in the context of experiments like mine. However, in the wider context of examining the variety of potential nonlinearities in the auditory processing of speech signals, the possibility I raise might be worth considering in future research.

A second, related possibility for why listeners might have tracked *F*n over F_0 is that *F*n was again the more salient feature, not in absolute terms, but rather in the degree to which shifts in *F*n affected assessments of size, masculinity and attractiveness compared to shifts in $F_{0.}$ In other words, while my shifts in *F*n and F_0 were equally *discriminable,* they may not have been equally salient in terms of the proportional difference they represented in each rating dimension. For example, because there is a two-fold difference in voice pitch between men and women but a difference of only about 15% in voice formants, a 6% change in voice *F*n might have a much greater effect on perceptions of the relative masculinity (or femininity) of a given voice than would the same 6% difference in voice F_0. Put differently, it requires only three such 6% shifts to move a voice that is canonically female in its *F*n profile to being canonically male, while it would require fifteen such 6% shifts to change a voice that is canonically female in its F_0 profile to being canonically male. Hence, the 6% shifts in *F*n and F_0 used in the current experiment might have represented a much greater proportion of the differences between males and females in *F*n compared to F_0. Again, it is unclear how this gender-based schema might favour *F*n over F_0 in all three social assessments (e.g., attractiveness) and requires additional investigation of the extent to which different voice-based social assessments are grounded in or biased by the perceived sex of the speaker.

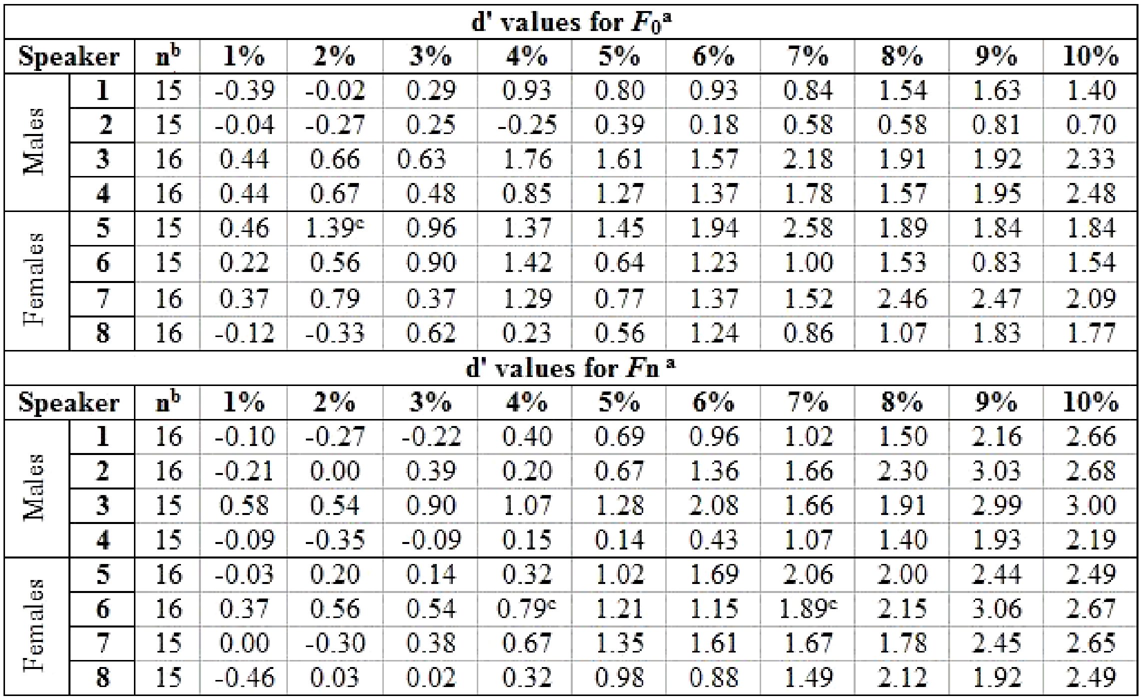
Third, listeners might have tracked *F*n over F_0 because the former is, in fact, a more reliable cue to the social dimensions being evaluated. Certainly, this may be the case for assessments of body-size where *F*n is a dependable acoustic correlate of VTL, as proposed by Fitch. Subsequent empirical work has generally confirmed Fitch’s proposal and documented a reliable three-way correlation between body-size, vocal-tract length and formant frequencies in several mammalian species (Fitch, 2007; Fitch & Giedd, 1999; Reby & McComb, 2003; Riede & Fitch, 1999). Studies of these relationships in humans have also generally reported similar correlations between body-size and either vocal-tract length or formant frequencies in both male (Evans *et al.*, 2006; Fitch & Giedd, 1999; González, 2004; Griesbach, 1999; Rendall *et al.*, 2005) and female speakers (Collins & Missing, 2003; González, 2004; Griesbach, 1999), though a small number have failed to find correlations for one or the other sex (Bruckert *et al.*, 2006; Rendall *et al.*, 2005; van Dommelen & Moxness, 1995). In contrast, research is consistent in finding that voice pitch is *not* strongly correlated with body-size within adults of either sex (Collins & Missing, 2003; Evans *et al.*, 2006; González, 2004; Hollien, Green, & Massey, 1994; Hollien & Jackson, 1972; Künzel, 1989; Majewski, Hollien, & Zalewski, 1972; Rendall *et al.*, 2005; van Dommelen & Moxness, 1995). Hence, listeners in our study might have tracked Fn over F_0 in their assessments of speaker size because formants are in fact the more reliable cue, at least within sexes.

Although it is sensible to assume that listeners used Fn as a cue to size because of its reliable correlations with height, it is not so apparent why listeners weighed Fn cues most heavily in assessments of masculinity (or femininity) and attractiveness. There is some evidence to suggest that testosterone increases somatic tissue development (Notelovitz, 2002), thus predicting VTL (Fitch and Giedd, 1999), and thus that Fn can in fact predict testosterone levels in men (Bruckert *et al.*, 2006). Otherwise, masculinity and femininity are closely associated (if not semantically interchangeable) with the differences in biological sex, that as noted earlier is most conspicuously cued by the two-fold difference in voice F_0 between men and women. Furthermore, in as much as assessments of attractiveness are rooted in assessments of maleness and femaleness (masculinity and femininity) it would seem that attractiveness ratings might also be rooted in the more conspicuous F_0 difference between men and women. Hence, the apparent prioritization of *F*n cues in the latter assessments, in addition to the assessment of size, is puzzling. The small possibility that *F*n may, in fact, be the more honest cue, not only to size but also to masculinity and attractiveness, will be discussed in more detail in Chapter Five.

At the same time, it is possible that the prioritization of *F*n cues in assessments of masculinity/femininity and attractiveness simply followed their prioritization as more reliable cues in size assessments and did so because of the inherent size-related associations among all three rating dimensions. Men and women differ consistently in size; hence, notions of masculinity and femininity, and the notions of attractiveness that might flow from them, could be commonly anchored by basic perceptions of size, with "larger voices" epitomizing masculinity and attractiveness in males and "smaller voices" epitomizing femininity and attractiveness in females.

If true, an important collorary is that size, masculinity/femininity and attractiveness might not actually represent independent rating dimensions in voice-based social judgments, even though investigators typically analyze them independently and generally assume that listeners can evaluate them independently. Instead, they might be intrinsically linked to one another and tied to basic perceptions of size that can be a relevant dimension of perceptual evaluation not just in the perception of human voices but in the perception and adaptive responding to many other environmental sounds (for further discussion of this possibility, see Grassi, 2005; Rendall *et al.*, 2007). This fourth and final potential explanation for the apparent prioritization of Fn over F_0 will be addressed in the following Chapter, where I test the degree to which ratings of size, masculinity and attractiveness overlap, or can be independent, in voice-based social judgments.

Table 3.1 Summary of *d'* scores for F_0 and *F*n frequency discrimination for each speaker.



d' values for F_0[a]												
Speaker		**n**[b]	**1%**	**2%**	**3%**	**4%**	**5%**	**6%**	**7%**	**8%**	**9%**	**10%**
Males	**1**	15	-0.39	-0.02	0.29	0.93	0.80	0.93	0.84	1.54	1.63	1.40
	2	15	-0.04	-0.27	0.25	-0.25	0.39	0.18	0.58	0.58	0.81	0.70
	3	16	0.44	0.66	0.63	1.76	1.61	1.57	2.18	1.91	1.92	2.33
	4	16	0.44	0.67	0.48	0.85	1.27	1.37	1.78	1.57	1.95	2.48
Females	**5**	15	0.46	1.39[c]	0.96	1.37	1.45	1.94	2.58	1.89	1.84	1.84
	6	15	0.22	0.56	0.90	1.42	0.64	1.23	1.00	1.53	0.83	1.54
	7	16	0.37	0.79	0.37	1.29	0.77	1.37	1.52	2.46	2.47	2.09
	8	16	-0.12	-0.33	0.62	0.23	0.56	1.24	0.86	1.07	1.83	1.77
d' values for *F*n [a]												
Speaker		**n**[b]	**1%**	**2%**	**3%**	**4%**	**5%**	**6%**	**7%**	**8%**	**9%**	**10%**
Males	**1**	16	-0.10	-0.27	-0.22	0.40	0.69	0.96	1.02	1.50	2.16	2.66
	2	16	-0.21	0.00	0.39	0.20	0.67	1.36	1.66	2.30	3.03	2.68
	3	15	0.58	0.54	0.90	1.07	1.28	2.08	1.66	1.91	2.99	3.00
	4	15	-0.09	-0.35	-0.09	0.15	0.14	0.43	1.07	1.40	1.93	2.19
Females	**5**	16	-0.03	0.20	0.14	0.32	1.02	1.69	2.06	2.00	2.44	2.49
	6	16	0.37	0.56	0.54	0.79[c]	1.21	1.15	1.89[c]	2.15	3.06	2.67
	7	15	0.00	-0.30	0.38	0.67	1.35	1.61	1.67	1.78	2.45	2.65
	8	15	-0.46	0.03	0.02	0.32	0.98	0.88	1.49	2.12	1.92	2.49

a. *d'* scores were calculated for each incremental increase in frequency from 1 to 10% according to the formula: *d'* = Z(Hits) – Z(False Alarms).
b. *n* = number of listeners discriminating the corresponding speaker.
c. False alarm rate (FA) = 0; *d'* scores calculated by assigning FA of 1 (Stanislaw & Todorow, 1999).

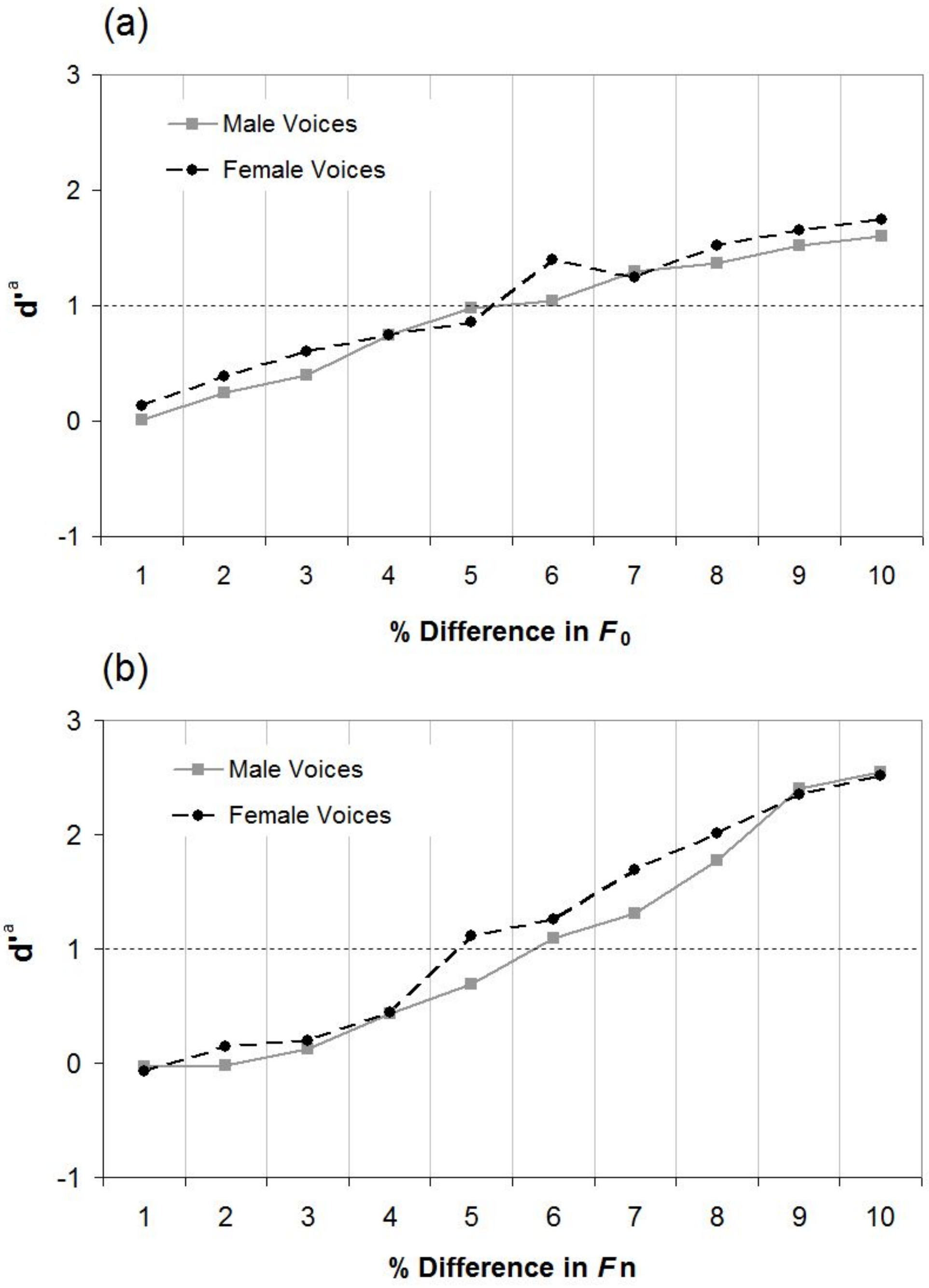


Figure 3.1. Mean *d'* scores^a for the discrimination of frequency differences in F_0 (a) or *F*n (b) as a function of speaker sex.

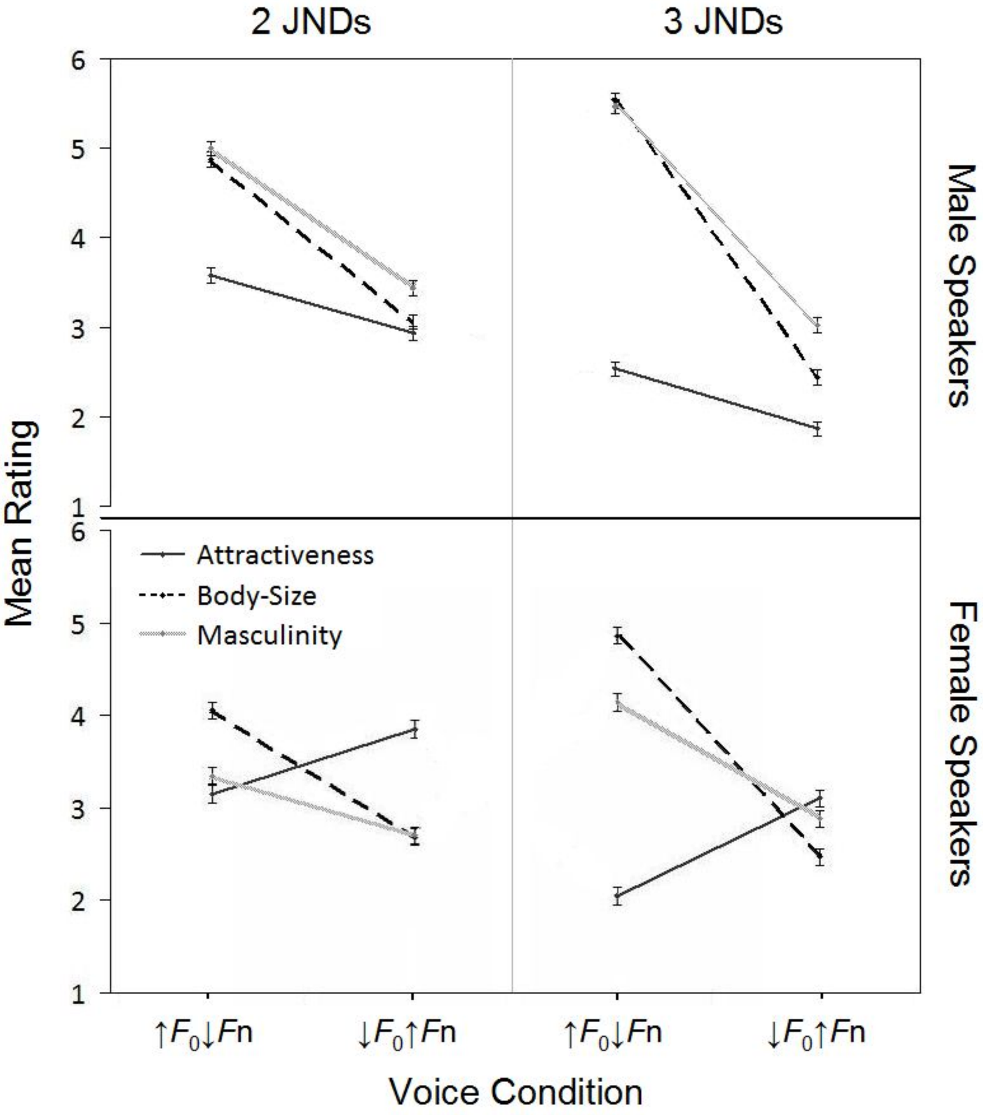


Figure 3.2. Mean (±SE) ratings of body-size, masculinity (femininity), and attractiveness for male and female speakers with manipulated high F_0 and low Fn (↑F_0 vs.↓Fn) compared to those with manipulated low F_0 and high Fn (↓F_0 vs.↑Fn). Separate panels are shown for manipulations of two and three JNDs. Ratings: 1 = small, feminine, or unattractive; 6 = large, masculine, or attractive. All effects shown are significant at $P < 0.05$.

CHAPTER FOUR

INTERDEPENDENCE IN VOICE-BASED BIOSOCIAL JUDGMENTS

4.1 INTRODUCTION

When making judgments about a speakers' size, masculinity or attractiveness from voice cues, listeners associate lower frequencies of either F_0 or Fn with largeness and masculinity and prefer them over higher frequencies in men but not women. If, however, listeners are provided with conflicting voice cues (i.e., ↓F_0 and ↑Fn or vice versa), assessments appear to be influenced more by cues from Fn than F_0. Although this outcome is sensible for assessments of body-size, where Fn is the stronger predictor of size, it is less clear why listeners track Fn in assessments of masculinity or attractiveness, where F_0 is the stronger predictor of sex and underlying hormonal state.

There may be several possible explanations for this curious outcome (discussed in Chapter 3 and further in Chapter 5). One possibility is that listeners' ratings of the different dimensions were not truly independent. Researchers might naturally treat the dimensions independently in their research designs and data analysis and they might assume that subjects can rate the dimensions independently. Nevertheless, these dimensions have relatively loose social definitions and implications and it is possible that peoples' conceptions of these dimensions are rather loose as a result. Furthermore, there is some natural overlap between size, masculinity and attractiveness generally. Taken together, it is possible that people's conceptions of these dimensions overlap appreciably and hence that their ratings of these dimensions in others based on voice (or other) stimuli are likewise correlated and not fully independent.

In this chapter, I attempt to test this possibility using principal component analysis (PCA). Using this method, it is possible to test the degree to which listeners' voice-based ratings of size, masculinity and attractiveness are in fact correlated or independent. Assuming that the design of my experiment did not allow for carry-over effects from ratings of one dimension to the next, any interdependence between listeners' ratings may thus be interpreted in light of the potential perceptual relations between them.

4.2 METHODS

4.2.1 Voice Stimuli

Selection, construction and acoustic analysis of voice stimuli is discussed in detail in the previous Chapter. In brief, natural voice stimuli were derived from 20 speakers (10 m, 10 f) and consisted of the words, "*bet, butt, bite, beat, book*", presented in their natural, baseline vocal frequencies. Each natural voice stimulus was then manipulated in four ways using PRAAT, by either raising F_0 while lowering Fn or by lowering F_0 while raising Fn and doing so by either 2 or 3 JNDs (see Chapter 3, Part I, for further discussion of JNDs); thus, yielding 80 unique, manipulated voice stimuli.

4.2.2 Participants

Thirty-two females and thirty-six males completed this experiment. All participants were recruited from the University of Lethbridge undergraduate community and self-reported as heterosexual.

4.2.3 Experimental Procedure

A detailed account of the experimental procedure can be found in the previous Chapter. Briefly, each participant (listener) was randomly assigned to one of two versions of the experiment. In Version A, voice stimuli and rating dimensions were completely randomized; in Version B, voice stimuli were blocked by rating dimension (either size, masculinity/femininity or attractiveness). Listeners in either version rated the same opposite-sex voices (10 natural and 40 manipulated) on each of three dimensions separately (size, masculinity/femininity, and attractiveness) for a total of 150 rating trials per participant. Ratings were performed using a 6-pt scale with no midpoint as described in Chapter Two.

4.2.4 Statistical Analysis

Rating data used for correlational analyses here are the same as those used in Chapter Three, Part II. Data were taken from this study in particular because of the large and representative sample of natural voice stimuli ($n = 20$); the large number of ratings obtained for each dimension ($n = 320 - 1440$); and the curious salience of Fn.

Subjects' response times (RT's) were collected for each rating trial and then analyzed using rmANOVA's to test for RT differences between sexes, stimulus types (i.e., natural vs. manipulated voices), dimensions, versions, and to control for any outliers. For all remaining analyses, listeners' data were analyzed separately for male and female voices and for natural and manipulated versions of them. Trials involving RT's less than 50-ms were removed prior to analysis. Rating data were collapsed across Versions A and B of the experiment because listeners' ratings did not differ significantly between them (see Results for Chapter 3, Part II).

I examined the degree to which ratings for a given dimension correlated with those for other dimensions using linear regression with an alpha of 0.05. I then conducted principal component analyses (PCA; NCSS v. 5.1, Hintze, 1989) with no rotation in order to examine the degree of interdependence between listeners' ratings of different dimensions and further to determine the degree of variance in ratings attributable to each set of dimensions or each principal component. To do this, ratings for each dimension (size, masculinity and attractiveness) were categorized as separate variables in the PCA and analyzed in terms of factor loadings. A conservative loading cutoff point of ±0.70 was used to include a variable in a component as loadings in excess of 0.71 are considered an excellent measure of a given factor (Jolliffe, 2002; Tabachnick & Fidell, 1989).

4.3 RESULTS

4.3.1 Correlations

Correlations between the three rating dimensions (size, masculinity and attractiveness) are reported in Table 4.1 for female's ratings of male voices and in Table 4.2 for male's ratings of female voices. In both male and female voices, all three dimensions were significantly correlated. Overall, higher correlations were observed between masculinity (or femininity) and body-size than any other two variables and these correlations were highest for manipulated voices.

In male voices in particular, all three rating dimensions correlated *positively* with one another; the strongest relationship observed was between ratings of size and masculinity. In female voices, ratings of attractiveness correlated *negatively* with those for size and masculinity, with the strongest relationship being between ratings of attractiveness and masculinity (femininity).

4.3.2 Factor Loadings and Principal Components

For females' ratings of natural male voices (baseline F_0 and Fn), the first factor was associated with masculinity and body-size and accounted for 47.99% (eigenvalue = 1.44) of the variance in ratings. The second factor corresponded to attractiveness and accounted for 28.47% (eigenvalue = 0.85) of the variance in ratings. For females' ratings of manipulated male voices, where baseline F_0 and Fn values were raised or lowered by 2 or 3 sex-specific JNDs, the first factor was again associated with masculinity and body-size (62.21%; eigenvalue = 1.87) while the second factor was associated with attractiveness (28.1%; eigenvalue = 0.84; Table 4.3).

For males' ratings of natural female voices, the first factor was associated with masculinity and attractiveness (55.43%; eigenvalue = 1.66) with the two factors negatively correlated with one another, while the second factor was associated with body-size (24.86%; eigenvalue = 0.75). Note that factor loadings for body-size were just short of the ±0.70 cutoff for Component 1 and just above it for Component 2 (see Table 4.4), making body-size ratings potentially more dependent on those for masculinity and attractiveness than my categorization might imply. For males' ratings of manipulated female voices, the first factor was associated with all three dimensions (64.98%; eigenvalue = 1.95; Table 4.4), with attractiveness ratings negatively correlated with both body-size and masculinity ratings.

4.3.3 Response Times

In general, subjects' RT's improved across trials (Figure 4.1). There were no significant differences between RT's for ratings of natural versus manipulated voices (m: $F_{1,\,31} = 1.83$, $P = 0.19$; f: $F_{1,\,35} = 0.73$, $P = 0.40$) or for ratings of size, masculinity and attractiveness (m: $F_{2,\,31} = 2.78$, $P = 0.07$; f: $F_{2,\,35} = 1.73$, $P = 0.18$).

Mean RT's for females' ratings of male voices ($M = 1.7$ s) were significantly faster than mean RT's for males' ratings of female voices ($M = 2.2$ s; $F_{1,\,66} = 5.73$, $P < 0.05$). Mean RT's were also faster in Version B than Version A of the experiment ($F_{1,\,66} = 5.21$, $P < 0.05$).

4.4 DISCUSSION

Results demonstrate significant correlations between all three rating dimensions. Principal component analyses confirm considerable interdependence between listeners' ratings of size, masculinity and attractiveness but some degree of separation for ratings of attractiveness in male voices and, to a lesser degree, ratings of body-size in female voices.

The directional relationships reported here between size, masculinity and attractiveness are intuitive and in line with previously reported findings from Chapters Two and Three showing female preferences for (or attraction to) masculine and large-sounding male voices and male preferences for feminine and small-sounding female voices. Other studies have also demonstrated similar relationships between perceived speaker size, masculinity or dominance, and attractiveness, where size and masculinity correlate positively with attractiveness in ratings of male voices and negatively in female voices (Berry, 1990; Collins, 2000; Feinberg *et al.*, 2005b, 2006, 2008a; Vukovic, Feinberg, Jones, DeBruine, Welling, Little & Smith, 2008). The question remaining then is *why* ratings of size, masculinity and attractiveness were so closely related to one another.

It might be argued that the high interdependence in ratings between dimensions reported here is merely a product of a within-subjects experimental design wherein each listener rated voices on all three dimensions. However, trials involving the same voice were often widely separated and all vocal stimuli were standardized for amplitude, length, speed, pronunciation and intonation (see Chapter 3), making it difficult for listeners to differentiate and track individual speakers whose voices were, in any case, manipulated across trials. As a result, it seems unlikely that listeners could recognize and remember which of 50 perceptually unique and randomly sorted vocal stimuli they had already heard and which they had not, and thus that past ratings would have any effect on future ratings.

Empirical substantiation that my results are not likely a product of a within-subjects design is obtained by comparing the PCA results reported above with other PCAs. If, for example, my results were due to within-subjects carry-over effects, then we might expect to see differences in the degree of interdependence between rating dimensions in the two versions of the experiment where stimuli were blocked by dimension (Version B) rather than presented completely at random (Version A). In fact, a comparison of PCAs conducted separately for each version of the experiment showed no sizeable differences between the factor loadings for either presentation schedule. Likewise, unpublished PCA results from a similar but *between-subjects* study conducted in our laboratory reveal relationships between dimensions akin to those reported here, where ratings of body-size and masculinity are highly codependent (Component 1) while ratings of attractiveness are independent from those for either size or masculinity (Component 2). These comparisons illustrate that a variety of experimental designs can elicit high correlations and codependences between listeners' ratings of different biosocial dimensions.

Thus, as a viable alternative interpretation, the interdependence of ratings between dimensions reported here may instead reflect a parallel interdependence in listeners' perceptions or categorizations of size, masculinity and attractiveness. In other words, although these three dimensions might be semantically distinct and are often treated as independent variables in psychological research, they may in fact overlap to some degree in listeners' perceptions. It is quite possible then that for listeners, rating a voice on size and masculinity is essentially the same task, where largeness and masculinity are closely related. Rather than assessing a voice on size (S) or masculinity (M) *per se,* the listener may in a sense be assessing the voice on size or masculinity using some combination of the two (e.g., S+M = M), hence perceiving these different dimensions holistically.

This perceptual holism may or may not reflect listeners' awareness of the true physical relationships that exist between body-size, masculinity or femininity, and attractiveness in the real world. If listeners were, however, either implicitly or explicitly mindful of the true correlations between these three dimensions then it may also be the case that listeners in this study used one dimension as a proxy for assessing another. Rather than relying on cues to masculinity (or size and masculinity, S+M) for assessing masculinity, then, they may instead have used size as a mediating factor for assessing masculinity (S→M) or attractiveness (S→A). Ultimately, it is difficult to say whether subjects in this study perceived voices holistically or used proxy's in their voice-based assessments and, if so, which dimension acted as the principle proxy. Nevertheless, there is little doubt that size, masculinity and attractiveness are predictive of one another in the real world and thus that such techniques may theoretically be successful (Bhasin, Storer, Berman, Yarasheski, Clevenger, Phillips, Lee, Bunnell & Casaburi, 1997**;** Boyapati, Shu, Gao, Dai, Yu, Cheng, Jin & Zheng, 2004; Collins, 2000; Collins & Missing, 2003; Jackson & McGill, 1996; Roney, Hanson, Durante, & Maestripieri, 2006; Thornhill & Gangestad, 1999; Yates *et al.,* 2004).

Separate PCAs of male and female voices revealed some divergence in ratings of attractiveness for male voices and ratings of body-size for female voices (see Tables 4.3 and 4.4) from ratings of the other two dimensions. Although all three dimensions correlate to some degree in the real world, there may be sensible explanations for sex-specific divergences in listeners' ratings. Firstly, male attractiveness has been found to be influenced by factors unrelated to a man's size, masculinity or dominance. Whereas males always prefer femininity in female features, research shows that females' voice- and face-based attributions of male attractiveness and masculinity can vary with their menstrual cycle (Feinberg *et al.,* 2006; Johnston, Hagel, Franklin, Fink, & Grammer, 2001; Puts, 2005) and that sometimes (e.g., for long-term versus short-term relationships) women prefer more feminine men (Penton-Voak & Perrett, 2000). These extenuating factors might be enough to account for the minor independence in attractiveness ratings from those of size and masculinity in male voices.

Secondly, the slight divergence in men's ratings of body-size from those of masculinity (femininity) and attractiveness in female voices might reflect a similar bifurcation in men's body preferences for women. Studies have shown that men's body preferences can vary within and between cultures (Greenberg & LaPorte, 1996; Swami & Tovee, 2009; Terry & Vasey, 1999) and may be more heavily influenced by body-shape than size (Singh, 1994).

As an important aside, it must be noted that concepts or percepts of size, masculinity and attractiveness need not overlap for subjects' voice-based assessments of these dimensions to be closely correlated, as they were here. This is because vocal features (F_0 and Fn) have been shown to cue all three dimensions independently but in similar ways, readily explaining the covariation between their ratings. Thus, the positive relationships documented here between ratings of size, masculinity and attractiveness in male voices (see Table 4.1) and the negative relationship documented between attractiveness and masculinity or size in female voices (see Table 4.2) may be incidental and may simply reflect the directional effects that vocal cues have on ratings of these dimensions and how those vary in male and female voices (i.e., low Fn/ high F_0 elicits high ratings of size, masculinity and attractiveness in male voices {↑S, ↑M, ↑A} and high ratings of size, masculinity but low ratings of attractiveness in female voices {↑S, ↑M, ↓A}). This possibility is discussed in more detail in the next Chapter, where I argue for the potential utility of Fn in predicting speaker masculinity and attractiveness, in addition to size.

Analysis of subject's RT's showed some predictable patterns but also highlighted further differences between the sexes in my sample. Naturally, RT's improved over the course of the experiment as subjects became more familiar with the rating task and were faster for Version B than Version A as the former was a less cognitively demanding task (i.e., the rating dimension did not vary nearly as frequently). Furthermore, listeners showed similar efficiency in judging manipulated versus natural voices and similarly in judging between size, masculinity and attractiveness. Perhaps less predictably, however, I found that females were faster at rating males' voices than males were at rating females' voices. This outcome may be due to some intrinsic difference between the rating regimes of male and female *subjects*, such as systematic sex differences in speed-accuracy trade-offs (Der & Deary, 2006; Reimers & Maylor, 2006). Alternatively, there may be some difference in the salience or ease of processing between male and female *voices*. For instance, Lattner, Meyer and Friederici (2005) report differences in fMRI brain activation in response to female and male voices but report no effect of listener's sex.

Whereas I found predictable differences in the factor loadings of male and female voices, there were no differences in the loadings between natural and manipulated voices within either sex (see Tables 4.3 and 4.4). This suggests that listeners' rating patterns or general algorithms for assessing size, sex and attractiveness in men and women were affected by the sex of the speaker but not by the sizeable manipulations performed on baseline vocal frequencies. In other words, listeners used the same rules for assessing voices with naturally-occurring F_0 and Fn as for assessing those with computer-manipulated vocal frequencies. This finding implies that the systematically-controlled, computer-manipulated vocal stimuli frequently used in psychoacoustic research do exhibit some external validity in their effects on listener's ratings.

The findings reported here yield additional practical implications for future psychological and psychoacoustic research. Future studies should be careful to control for sex effects of vocal stimuli and, further, possible interactions with the sex of the listener. Most important, researchers should be cautious in interpreting subjects' ratings of seemingly independent social features or dimensions such as body-size, masculinity or sex, and attractiveness. Although subjects' perceptions and interpretations of these dimensions may in actuality mirror their autonomous semantic categorizations, it is also feasible that these dimensions are perceived by subjects in a more combinatory or holistic way. To be sure, future perceptual research might benefit from a detailed analysis of any real or perceived relationships between a given set of rating dimensions.

Table 4.1. Correlations between rating dimensions for females' ratings of natural and manipulated male voices.

	Dimension		
Dimension	body-size	masculinity	attractiveness
	Natural Voices[a]		
body-size	$r = 1.00$		
masculinity	$r = 0.29$ $r^2 = 0.08$	$r = 1.00$	
attractiveness	$r = 0.16$ $r^2 = 0.03$	$r = 0.20$ $r^2 = 0.04$	$r = 1.00$
	Manipulated Voices[a]		
body-size	$r = 1.00$		
masculinity	$r = 0.71$ $r^2 = 0.50$	$r = 1.00$	
attractiveness	$r = 0.27$ $r^2 = 0.07$	$r = 0.26$ $r^2 = 0.07$	$r = 1.00$

a. r and r^2 values obtained from linear regression tests. All correlations are significant at $P < 0.05$.

Table 4.2. Correlations between ratings dimensions for males' ratings of natural and manipulated female voices.

	Dimension		
Dimension	body-size	masculinity	attractiveness
Natural Voices[a]			
body-size	$r = 1.00$		
masculinity	$r = 0.30$ $r^2 = 0.09$	$r = 1.00$	
attractiveness	$r = -0.28$ $r^2 = 0.08$	$r = -0.41$ $r^2 = 0.17$	$r = 1.00$
Manipulated Voices[a]			
body-size	$r = 1.00$		
masculinity	$r = 0.54$ $r^2 = 0.29$	$r = 1.00$	
attractiveness	$r = -0.43$ $r^2 = 0.18$	$r = -0.45$ $r^2 = 0.20$	$r = 1.00$

a. r and r^2 values obtained from linear regression tests. All correlations are significant at $P < 0.05$.

Table 4.3. Unrotated factor loadings for PCA of females' ratings of natural and manipulated male voices.

Dimension	Component 1[a]	Component 2[a]
Natural Voices		
body-size	**-0.713**	-0.451
masculinity/femininity	**-0.750**	-0.205
attractiveness	-0.608	**0.781**
Manipulated Voices		
body-size	**-0.890**	-0.247
masculinity/femininity	**-0.887**	-0.262
attractiveness	-0.535	**0.845**

a. ±0.70 was the cutoff point used for including a dimension in a Component and factor loadings meeting this criterion are bolded.

Table 4.4. Unrotated factor loadings for PCA of males' ratings of natural and manipulated female voices.

Dimension	**Component 1**[a]	**Component 2**[a]
Natural Voices		
body-size	-0.680	**0.731**
masculinity/femininity	**-0.782**	-0.268
attractiveness	**0.768**	0.375
Manipulated Voices		
body-size	**-0.820**	0.348
masculinity/femininity	**-0.831**	0.247
attractiveness	**0.765**	0.641

a. ±0.70 was the cutoff point used for including a dimension in a Component and factor loadings meeting this criterion are bolded.

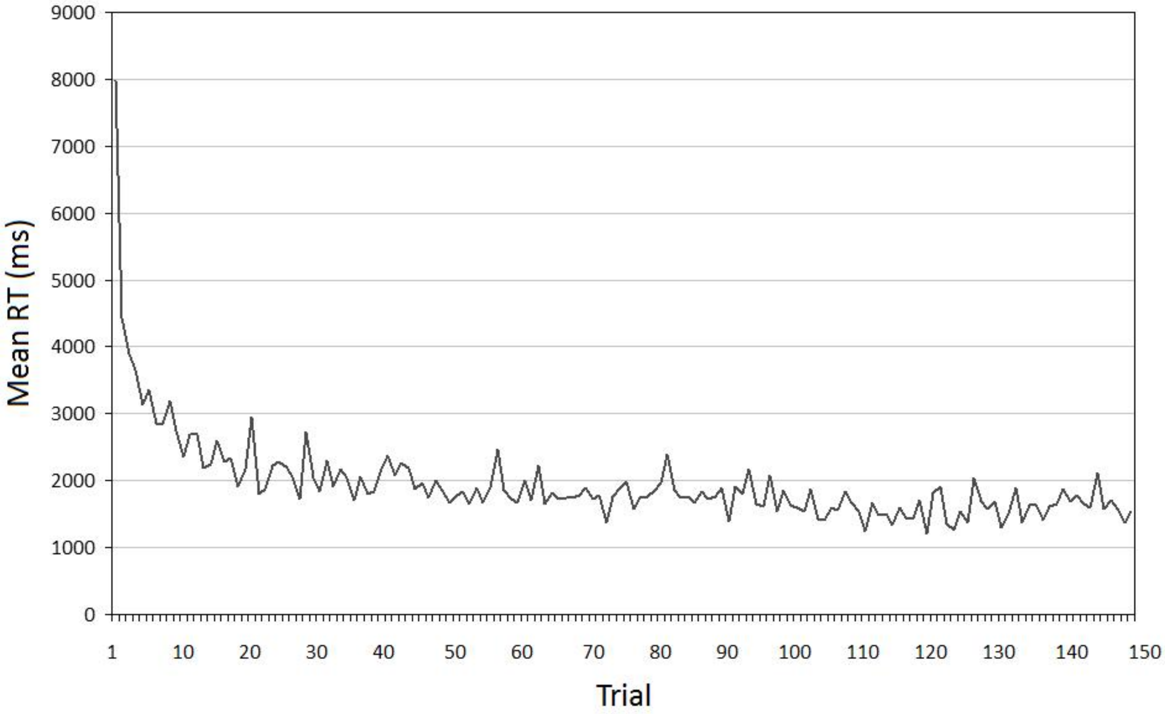


Figure 4.1. Mean RT's for ratings of size, masculinity and attractiveness in opposite-sex voices collapsed across *n* = 68 male and female subjects.

CHAPTER FIVE

SUMMARY

Research on vocal cueing in humans corroborates analogous work in non-human communication showing that vocalizations can cue numerous qualities of social and evolutionary significance. In addition to showing that vocal features correlate with the physical, physiological and hormonal states of the signaler (Cowlishaw, 1996; Galeotti *et al.,* 1997; Harris *et al.,* 2006) which might have fitness implications (Eriksson & Wallin, 1986), animal studies have further shown that listeners can be sensitive to these voice-based relationships (Charlton, Reby & McComb, 2007; Charlton, Zhihe & Snyder, 2010). In other words, listeners actively attend to vocal cues, and the underlying acoustic features of vocalizations (e.g., F_0 and Fn) can, in and of themselves, affect listeners' social perceptions and behaviour (Charlton *et al.,* 2007; Owren & Rendall, 1997, 2001; Rendall & Owren, 2010).

Vocal F_0 and Fn cues are thought to be rendered reliable, or 'honest', by the anatomical and hormonal constraints imposed on them (for review of vocal signals and constraints in animals see Fitch and Hauser, 2003). Thus, in addition to earlier work discerning the nature and extent of indexical information that can be obtained from the voice in general (Appleby & Redpath, 1997; Clutton-Brock & Albon, 1979), recent work in animal communication has adopted a source-filter framework to better understand which vocal features are driving the effects and why (Munhall and Byrne, 2007).

Using this framework, Fitch (1994) has argued that the resonances of the vocal-tract (i.e., *F*n) will more honestly cue body-size within-groups (e.g., species and sexes) than cues from the vocal source (i.e., F_0); this is because of *F*n's more intimate relationship to VTL (see Chapter One). Indeed, *F*n has been shown to predict body-size within species and sexes (mostly males) better than F_0. This effect has been demonstrated in colobus monkeys (Harris *et al.,* 2006), red deer (Reby & McComb, 2003), giant pandas (Charlton *et al.,* 2009), and humans (Collins & Missing, 2003; Evans *et al.,* 2006; González, 2004; Griesbach, 1999; Rendall *et al.,* 2005, 2007).

In subsequent human work, F_0 and *F*n have been shown to cue a variety of different speaker qualities including: speaker sex (Childers & Wu, 1991; Coleman, 1976) or masculinity/femininity (Bruckert *et al.,* 2006; Dabbs & Mallinger, 1999; Feinberg *et al.,* 2005a); strength or dominance (Puts, Gaulin,& Verdolini, 2006; Puts *et al.*, 2007; Sell *et al.,* 2010); identity (Bachorowski & Owren, 1999; Owren & Cardillo, 2006); age (Feinberg, 2004; Hollien *et al.,* 1994); body-size (Bruckert *et al.,* 2006; Dusan, 2005; Evans *et al.,* 2006; Fitch & Giedd, 1999; González, 2004; Rendall *et al.*, 2005, 2007; Smith *et al.*, 2005); health (Feinberg, 2004); fertility or reproductive value (Abitol *et al.,* 1999; Alonso *et al.,* 2002; Apicella *et al.,* 2007; Hughes *et al.,* 2004); emotional state or personality (de Gelder and Vroomen, 2000; Zuckerman *et al.,* 1989); and attractiveness (Berry, 1990; Bruckert *et al.,* 2006; Collins, 2000; Collins & Missing, 2003; Feinberg *et al.,* 2005b). Of these many traits, my research has focused on three in particular: size, masculinity and attractiveness. The rationale for this is that these traits are, arguably, most integral to human social interaction and are especially relevant in assessments of mates or rivals. Focusing on these three traits, and body-size in particular, has also allowed for cross-species comparisons and corroborations of the species-wide efficacy of vocal cues.

In testing the independent and joint effects that vocal cues of F_0 and *F*n have on listeners' voice-based assessments of these traits, I was able to replicate previous work showing that lower vocal frequencies in general are associated with greater masculinity and a larger body-size in both male and female speakers and attractiveness only in male speakers. Interestingly, I was also able to show that listeners' voice-based assessments of masculinity and size are quite general in nature and could be elicited using either natural or manipulated and either male or female voices.

I was further able to demonstrate that, when placed in conflict using JND units, F_0 and *F*n cues can have differential effects. Using the source-filter framework and theory of honest signaling for my own hypotheses regarding the relative effects of *F*n versus F_0, I predicted that listeners would track *F*n in assessments of size but F_0 in assessments of masculinity and attractiveness. These predictions were based on disparate accounts of honest signaling, where *F*n correlates reliably with size (*sensu* Fitch, 2000) while F_0 correlates with androgen levels (Dabbs & Mallinger, 1999). Results show that in fact listeners tracked *F*n over F_0 in all three assessments. This outcome confirms my prediction for assessments of body-size and supports analogous findings from the non-human literature suggesting that *F*n is the more reliable correlate (Fitch, 1997) as well as the more salient cue (Charlton, Reby & McComb, 2008) to body-size in vocalizing animals. However, this outcome challenges my prediction for F_0's role in assessments of masculinity and attractiveness. In the following section, I will discuss these experimental findings from two perspectives: influences of the voice and influences of the listener.

5.1 UNRESOLVED ISSUES

5.1.1 Formant Frequency as a Reliable Cue to Masculinity and Attractiveness

As reviewed in previous Chapters, listeners might have tracked *F*n over F_0 because *F*n was more audible when placed in opposition to F_0 (see Chapter 3, Part II), or perhaps as a result of overlapping concepts of size, masculinity and attractiveness (see Chapter 4 and proceeding section 5.1.2). Nonetheless, it is also possible that, contrary to my initial instinct, vocal *F*n may more reliably cue masculinity and attractiveness than F_0. In such a scenario, listeners in my study may have in fact been correctly tracking the more honest cue, not only for assessments of size, but also for assessments of masculinity/femininity and attractiveness.

While this theory is plausible, empirical substantiation for honesty in *F*n cues to masculinity and attractiveness is sparse and, further, complicated by the vague nature of individual and social definitions of “masculinity” and “attractiveness”. This presents a difficulty then in quantifying these social constructs accurately and in comparing them with measures of “dominance” and “preference” in non-human animals. Even so, some animal and human studies have shown stronger effects of *F*n than F_0 on signal-receiver behaviours that can loosely be interpreted as cues to masculinity (dominance) and attractiveness (preference), or at the least, cues to something other than just body-size.

Thus, Reby and McComb (2003) found that Fn, but not F_0, was able to predict reproductive success in red deer stags. Although this is not a direct measure of male attractiveness or dominance *per se*, it is likely to reflect some success on the male's part in securing mates and competing against other males. Others report more direct evidence. Charlton, Reby and McComb (2007, 2008) have demonstrated female preferences for (attraction to) red deer stags with lower formants with no effect of F_0 on females' responses to males. In the human literature, Childers and Wu (1991) report similar recognition of gender (masculinity/femininity) from $F2$ as from F_0. Others, using JNDs for manipulations in F_0 and Fn, have found stronger *relative* effects for Fn over F_0 in women's assessments of men's dominance (Puts *et al.*, 2007) and attractiveness (Puts, Barndt, Welling, Dawood, & Burris, 2010). A more focused research paradigm is necessary to properly test the relative utility of Fn over F_0 cues; however, these studies provide preliminary support for reliable cueing of masculinity and attractiveness through Fn.

5.1.2 Interdependence in Perceived Size, Masculinity and Attractiveness

To understand the independent and interactive effects of vocal F_0 and Fn on listeners' assessments of size, masculinity and attractiveness, we must also understand the nature of the dimensions themselves, both in terms of how they are measured in speakers as well as how they are defined by listeners.

The methodologies used for measuring these biosocial traits in speakers are relatively established but still show appreciable cross-study variability. For example, speaker masculinity and femininity are regularly quantified using assays of testosterone and estrogen, respectively (Bruckert *et al.,* 2006), but may also be quantified using androgen-mediated facial-metrics (Feinberg *et al.,* 2005a; Pound, Penton-Voak, & Surridge, 2009). Likewise, speaker attractiveness can be measured either physically (e.g., facial symmetry, adiposity, or complexion; Coetzee, Perrett, & Stephen, 2009; Fink, Grammer, & Matts, 2006) or by using averaged or relative preference ratings (Hughes *et al.,* 2004; Saxton *et al.,* 2006). Body-size can also be measured in various ways, using speaker height, weight, BMI, or body-shape (e.g., SHR or various circumferences, Evans *et al.,* 2006).

To complicate things further, we cannot be sure that the measures we use to quantify these traits are the same measures that listeners are attending to when they rate them. As noted earlier, precise definitions of masculinity and attractiveness, and even body-size, are difficult to ascertain. Ratings of masculinity might reflect variable ideas of sex/gender, strength, manliness, dominance, or even sexual orientation. Likewise, ratings of size *(<small>* versus *<large>*) may reflect perceptions of speaker height, weight, shape or a combination of the three. Indeed, there has been some debate in the academic literature regarding how people really perceive of and define masculinity (Thompson, Pleck and Ferrera, 1992) and size (Grassi, 2005).

What are we really measuring then? This question was addressed in Chapter Four where I tested how listeners' ratings of size, masculinity and attractiveness correlated or were independent of one another. I found that, perhaps not surprisingly, ratings for all three dimensions were interdependent to some degree. One possible explanation for this finding is that listeners' perceptions of these dimensions overlap extensively, such that, studies like my own may not be collecting entirely independent ratings of speaker size, masculinity and attractiveness but rather ratings that are anchored by one dimension or that better represent some holistic interpretation of the speaker's general quality.

Future studies may continue to address the ambiguity and potential holism in listeners' voice-based ratings of these dimensions by increasing the number (while narrowing the broadness) of dimensions being rated. For instance, rather than having subjects rate perceived “size” and “masculinity” directly, Collins (2000) asked her subjects to rate men's voices on weight, height, body-type, muscularity and chest hairiness. A principal component analysis might then allow researchers to discern the degree to which these sub-categories (e.g., chest hairiness) reflect the umbrella category to which they are assumed to belong (e.g., masculinity), and furthermore, to what degree sub-categories of different umbrella categories relate to one another. Although Collins did not perform a PCA on subjects' ratings, she did find high consensus among judges and close correlations between their assessments of attractiveness, age, hairiness, height, weight and muscularity. Her findings thus support the notion that listeners' conceptions of masculinity, size and attractiveness are in some ways appreciably overlapped.

5.1.3. Methodological Shortcomings

Traditionally, studies on human vocal signaling have employed a within-subjects experimental design when testing the effects of voice cues on ratings of different social and physical dimensions (Bruckert *et al.*, 2006; Collins & Missing, 2003; Feinberg *et al.*, 2006). Hence, each subject typically rates vocal stimuli on each relevant dimension, albeit, usually not simultaneously. Although in perceptual voice research this conventional method is both efficient and, due to randomization techniques and a large number of unique vocal stimuli, unlikely to affect experimental findings, there remains the possibility that listeners' ratings of one voice or dimension influence their ratings of others. Thus, in my experiments, I cannot be entirely certain that the correlations reported between ratings of size, masculinity, and attractiveness are not a product of subject carry-over effects. Future studies may address this issue by using a between-subjects design.

Another methodological tradition in research on vocal cueing is the use of either natural voice stimuli or stimuli that have been manipulated in absolute rather than perceptually equivalent (e.g., JND) units. Thus, in using JNDs for my stimulus manipulations, I have addressed the longstanding problem of perceptual inequality between F_0 and Fn. However, the specific subset of JNDs that I used introduces another potential shortcoming of my research, where, rather than having used sex-specific JNDs I might have more appropriately used speaker-specific JNDs. Indeed, my analysis of discrimination thresholds for F_0 and Fn revealed variability in listeners' ability to discriminate shifts in the voices of different speakers, in addition to different sexes. Thus, JNDs varied in the order of ±0-1% between male and female voices but ±2% across different speakers within a single sex (see Table 3.1). In my study, then, some voices were likely more discriminable than others (i.e., where sex-specific JNDs > speaker-specific JNDs) and may have influenced the outcomes observed. Future research should address this problem by, first, quantifying discrimination thresholds for each *speaker* in the sample, and, second, using speaker-specific JNDs in vocal manipulations.

A final, related issue concerns a lack of control in discrimination thresholds as a function of vowel sound. As illustrated in Table 5.1 (and in Kewley-Port & Watson, 1994; Smith *et al.*, 2005), baseline F_0 and Fn values can vary across vowel sounds and in predictable ways that reflect the shape or position of the mouth and tongue during pronunciation (Ladefoged, 2001). Thus, words like “boot” and “beat” have naturally lower baseline resonances or formant frequencies than words like “bet” and “butt”. Similar to the problem of variable JNDs across speakers then, discrimination thresholds for F_0 and Fn are also likely to vary across different vowel sounds within speakers (Kewley-Port, 1995). And, although the stimuli that I used purposefully contained four to five bVt words each with different vowel sounds to account for this variability, using a generic JND for manipulations of each of these different sounds within a given speaker suggests that some words may have been more discriminable than others.

Although it may be unrealistic to suppose that researchers have the time and resources to conduct separate JND analyses for each experiment, vocal feature, speaker, and sound produced by a speaker, it may nevertheless be a necessary preliminary step to producing proper vocal manipulations of perceptual equivalence. Alternatively, it may be possible to generate a one-time database of such a nature in order to discern any universals in F_0 and Fn discrimination. The discrimination differences between speakers in my sample may, for example, be predictable if they reflect some currently unidentified, underlying differences between speakers (e.g., nationality in addition to dialect, age, and sex). This effort would, of course, require collecting discrimination data for a great variety of speakers from different “groups” (e.g., sexes, ethnicities) and for a great variety of different sounds (e.g., vowels, consonants, whole words). Once such a database has been established, however, any consistent, between-group patterns in JNDs could be applied as appropriate to vocal manipulations in future experiments.

5.2 FUTURE DIRECTIONS

In addition to resolving some of the issues reviewed above, future research on vocal signaling might benefit from taking a more comprehensive approach to studying the effects of vocal cues on listener's biosocial assessments. We may begin to do so by using vocal stimuli that more accurately represent the natural human voice. We must, of course, first understand which vocal features affect listeners' judgments by systematically and independently manipulating features like F_0 and Fn or other vocal features such as jitter (deviations in vocal pulse-rate or pitch) and shimmer (deviations in voice amplitude; for experiments analyzing vocal jitter and shimmer see Baumann & Belin, 2010; Feinberg, 2004; Kreiman, Gerratt, & Gabelman, 2002). However, once we have a comprehensive understanding of its respective components and their effects, we may then move forward and study the voice as a whole. In doing so, we can test how the many different features of the voice interact, as their interactions may reveal novel, emergent effects on listeners' assessments. We can test how these features interact with other vocal qualities such as accent, intonation, speed of speech, and dialect. We can also test whether listeners perceive different vocal features (e.g., F_0 and Fn) independently or, akin to their overlapping perceptions of social dimensions, more holistically as some end product that reflects the general personality or quality of an individual's voice.

A more comprehensive analysis of vocal cues would further require a better understanding of how listeners' ratings or judgments reflect their cognition. In other words, we must also strive to appreciate what it is that our rating data truly represent and to understand the interactions between listeners' conceptions of speaker dimensions (like masculinity and size) and how those might bias their apparently independent ratings of them.

Last, a true comprehensive understanding of the effects of voice would require integration with other modes of human communication and perception (i.e., cues from the face and body). After all, in everyday interaction, the voice is often perceived in unison with these visual modes of cueing and, indeed, recent work has found parallels in individuals' vocal, facial and bodily dimensions. In two separate studies, persons receiving high vocal attractiveness scores also received high facial attractiveness scores (Collins & Missing, 2003; Saxton *et al.*, 2006). Similarly, researchers have documented stronger preferences for faces of women with high-pitched voices over faces of women with low-pitched voices (Feinberg *et al.*, 2005a), as well as preferences for voices of persons with more symmetrical bodies (Hughes *et al.*, 2002, 2008). Finally, researchers have also found correlations between peoples' preferences for vocal and facial masculinity (Feinberg *et al.*, 2008b). These studies have laid the groundwork for a comprehensive approach to the study of voice and face and, in doing so, have demonstrated that voice cues may in fact provide the same indexical information as face and body cues, simply strengthening the overall signal (Feinberg, 2008). Equally as important, they suggest that vocal cues follow the same laws of attraction, so to speak, as the face and body, and that neither visual nor acoustic cues can be understood fully when studied in complete isolation of one another.

5.3 RESEARCH SIGNIFICANCE

In so far as I have been able to demonstrate the efficacy of vocal cues through their effects on voice-based assessments of speakers, my research provides evidence for vocal cueing in humans and, likewise, for perceptual mechanisms in listeners that allow for meaningful interpretation of vocal cues. My research shows that listeners are sensitive to natural and synthetic variations in F_0 and Fn within and between sexes and that these variations have consistent and largely predictable effects on listeners' assessments of speaker size, masculinity and attractiveness. These findings substantiate earlier work in non-human animals; namely, work demonstrating female preferences for males with low vocal frequencies.

It is widely appreciated that low F_0 and Fn can cue or signal large size, dominance, and masculinity in men and other male animals and that females prefer these vocal features to higher frequencies. My research on vocal cues in women thus contributes to the general literature on vocal signalling by demonstrating that, while low frequencies function similarly to cue largeness and masculinity in female speakers, they do not function similarly to cue attractiveness. Instead, men appear to prefer women with high F_0 and Fn and this finding is sensible from an evolutionary perspective where these vocal features in women correlate with physiological measures of fertility and femininity (Abitol *et al.*, 1999; Bryant & Haselton, 2009; Hughes & Gallup, 2008).

Perhaps most important, the experimental design employed here to test the relative effects of F_0 and Fn is novel. As a result, I have been able to demonstrate for the first time that when placed in conflict with one another using perceptually equivalent frequency manipulations, cues from *F*n play a larger role in listeners' assessments of size, masculinity and attractiveness than cues from F_0. These findings support evolutionary theories of reliable or honest vocal cues to body-size and offer an interesting perspective on the role of formants in listeners' assessments of masculinity and attractiveness. To be precise, this research suggests that there may be some honesty in *F*n cues to these biosocial dimensions. Alternatively, I also provide compelling support for interdependence in listeners' social interpretations of size, masculinity and attractiveness that may have important methodological implications for future work in human psychoacoustics.

Table 5.1. Summary of F_0 and *F*1 – *F*4 values for eight different vowels in a bVt context.

bVt word[b]	**Vowel**	***n* (voices)**[a]	**F_0**[c]	***F*1**[c]	***F*2**	***F*3**	***F*4**
Male Voices							
bet	/ɛ/	13	98	572	1706	2601	3729
bite	/aɪ/	13	107	499	1787	2509	3505
bait	/e/	8	110	353	2179	2780	3597
beat	/i/	16	109	247	2305	3109	3664
boat	/Θ/	4	132	392	921	2417	3262
butt	/ɘ/	11	93	561	1289	2464	3490
boot	/u/	4	116	277	1306	2322	3382
book	/ʊ/	15	106	462	1131	2406	3405
Female Voices							
bet	/ɛ/	13	188	731	2011	3010	4169
bite	/aɪ/	11	196	667	1925	2854	3856
bait	/e/	8	150	369	2374	2966	3856
beat	/i/	17	189	309	2722	3325	4237
boat	/Θ/	6	216	509	1184	2848	3983
butt	/ɘ/	11	171	685	1434	2681	3763
boot	/u/	4	201	458	1603	2735	3822
book	/ʊ/	15	171	549	1237	2735	3774

a. Voices include only the natural (baseline F_0 and *F*n) voice stimuli and those used in all three experiments of this thesis (Chapters 2 and 3). Due to variation in stimulus word-sets, either between speakers or experiments, some words include measurements from a larger set of voices than others.

b. Due to a small sample of voices (i.e., $n = 1$), bVt words "bit" (/ɪ/) and "bat" (/æ/) have been omitted from the table.

c. Vocal frequencies (F_0 and *F*n) are presented in hertz (Hz).

REFERENCES

Abitol, J., Abitol, P., and Abitol, B. 1999. Sex hormones and the female voice. *Journal of Voice*, 13, 424-446.

Alexander, R.D., Hoogland, J.L., Howard, R.D., Noonan, K.M., and Sherman, P.W. 1979. Sexual dimorphism and breeding systems in pinnipeds, ungulates, primates and humans. In: *Evolutionary Biology and Human Social Behaviour: An Anthropological Perspective*. (Ed. by N.A. Chagnon and W. Irons), pp. 402-435. North Scituate, MA: Duxbury Press.

Alonso, L.C., and Rosenfield, R.L. 2002. Oestrogens and puberty. *Best Practice & Research Clinical Endocrinology & Metabolism,* 16, 13-30.

Apicella, C.L., Feinberg, D.R., and Marlowe, F.W. 2007. Voice pitch predicts reproductive success in male hunter-gatherers. *Biology Letters,* 3, 682-684.

Appleby, B.M., and Redpath, S.M. 1997. Indicators of male quality in the hoots of tawny owls (*Strix aluco*)*. Journal of Raptor Research,* 31, 65-70.

Bachorowski, J.-A., and Owren, M.J. 1999. Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *Journal of the Acoustic Society of America,* 106, 1054-1063.

Baken, R.J. 1987. *Clinical Measurement of Speech and Voice.* London: Taylor & Francis.

Baumann, O., and Belin, P. 2010. Perceptual scaling of voice identity: Common dimensions for different vowels and speakers. *Psychological Research,* 74, 110-120.

Berry, D.S. 1990. Vocal attractiveness and vocal babyishness: Effects on stranger, self and friend impressions. *Journal of Nonverbal behaviour,* 14, 141-153.

Bhasin, S., Storer, T.W., Berman, N., Yarasheski, K.E., Clevenger, B., Phillips, J., Lee, W.P., Bunnell, T.J., Casaburi, R. 1997. Testosterone replacement increases fat-free mass and muscle size in hypogonadal men. *Journal of Clinical Endocrinology and Metabolism,* 82, 407–413.

Boersma, P.P.G. 2001. Pratt, a system for doing phonetics by computer. *Glot International*, 5, 341-345.

Boyapati, S.M., Shu, X.O., Gao, Y.-T., Dai, Q., Yu, H., Cheng, JR., Jin, F., and Zheng, W. 2004. Correlation of blood sex steroid hormones with body-size, body fat distribution, and other known risk factors for breast cancer in post-menopausal Chinese women. *Cancer Causes and Control*, 15, 305-311.

Bruckert, L., Liénard, J.-S., Lacrois, A., Kreutzer, M. and Leboucher, G. 2006. Women use voice parameters to assess men's characteristics. *Proceedings of the Royal Society: Biological Sciences*, 273, 83-89.

Bryant, G.A. and Haselton, M.G. 2009. Vocal cues of ovulation in human females. *Biology Letters,* 5, 12-15.

Charlton, B.D., Reby, D., and McComb, K. 2007. Female red deer prefer the roars of larger males. *Biology Letters,* 3, 382-385.

Charlton, B.D., Reby, D., and McComb, K. 2008. Effect of combined source ($F0$) and filter (formant) variation on red deer hind response to male roars. *Journal of the Acoustic Society of America,* 123, 2936-2943.

Charlton, B.D., Zhihe, Z., and Snyer, R.J. 2010. Giant pandas perceive and attend to formant frequency variation in male bleats. *Animal Behaviour,* 79. 1221-1227.

Childers, D., and Wu, K. 1991. Gender recognition from speech. Part II: Fine analysis. *Journal of the Acoustic Society of America,* 90, 1841-1856.

Clutton-Brock, T.H., and Albon, S.D. 1979. The roaring of red deer and the evolution of honest advertising. *Behaviour,* 69, 145 -170.

Coetzee, V., Perrett, D.I., and Stephen, I.D. 2009. Facial adiposity: A cue to health? *Perception,* 38, 1700-1711.

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

Collins, S.A. 2000. Men's voices and women's choices. *Animal Behaviour,* 60, 773-780.

Collins, S.A. and Missing, C. 2003. Vocal and visual attractiveness are related in women. *Animal Behaviour,* 65, 997-1004.

Cowlishaw, G. 1996. Sexual selection and information content in gibbon song bouts. *Ethology,* 102, 272–284.

Dabbs, J.M., Jr., and Mallinger, A. 1999. High testosterone levels predict low voice pitch among men. *Personality and Individual Differences,* 27, 801-804.

Darwin, C. 1871. *The Descent of Man, and Selection in Relation to Sex.* London: John Murray.

de Gelder, B., and Vroomen, J. 2000. The perception of emotions by ear and by eye. *Cognition and Emotion,* 14, 289-311.

Der, G., and Deary, I.J. 2006. Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging,* 21, 62-73.

Dusan, S. 2005. Estimation of speaker's height and vocal-tract length from speech signal. In: *Interspeech-2005,* 1989-1992.

Dusan, S., & Deng, L. 1999. Vocal-tract length normalization for acoustic-to-articulatory mapping using neural networks. *Journal of the Acoustic Society of America,* 106, Pt. 2, 2181.

Eriksson, D., and Wallin, L. 1986. Male bird song attracts females – a field experiment. *Behavioral Ecology & Sociobiology,* 19, 297-299.

Evans, S.N., Neave, N., and Wakelin, D. 2006. Relationship between vocal characteristics and body size and shape in human males: An evolutionary explanation for a deep male voice. *Biological Psychology,* 72, 160-163.

Evans, S.N., Neave, N., Wakelin, D., and Hamilton, C. 2008. The relationship between testosterone and vocal frequencies in human males. *Physiology and Behaviour,* 93, 783-788.

Fant, G. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton and Co.

Feinberg, D.R. 2004. Fundamental frequency perturbation indicates perceived health and age in male and female speakers. *Journal of the Acoustical Society of America,* 115, 2609.

Feinberg, D.R. 2008. Are human faces and voices ornaments signaling common underlying cues to mate value? *Evolutionary Anthropology,* 17, 112-118.

Feinberg, D.R., DeBruine, L.M., Benedict, C.J., and Perrett, D. 2008a. The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception,* 37, 615-623.

Feinberg, D.R., DeBruine, L.M., Jones, B.C., and Little, A.C. 2008b. Correlated preferences for men's facial and vocal masculinity. *Evolution and Human Behaviour,* 29, 233-241.

Feinberg, D.R., Jones, B.C., Law Smith, M.J., Moore, F.R., DeBruine, L.M., Cornwell, R.E., Hillier, S.G., and Perrett, D.I. 2006. Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behaviour,* 49, 215-222.

Feinberg, D.R., Jones, B.C., DeBruine, L.M., Moore, F.R., Law Smith, M.J., Cornwell, R.E., Tiddeman, B.P., Boothroyd, L.G., and Perrett, D.I. 2005a. The voice and face of woman: One ornament that signals quality? *Evolution and Human Behaviour,* 26, 398-408.

Feinberg, D.R., Jones, B.C., Little, A.C., Burt, D.M., and Perrett, D.I. 2005b. Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour,* 69, 561-568.

Fink, B., Grammer, K., and Matts, P.J. 2006. Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evolution and Human Behaviour,* 27, 433-442.

Fitch, W.T. 1994. *Vocal-tract length perception and the evolution of language*. Unpublished doctoral dissertation, Brown University, Providence.

Fitch, W.T. 1997. Vocal-tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustic Society of America,* 102, 1213-1222.

Fitch, W.T. 1999. Acoustic exaggeration of size in birds via tracheal elongation: Comparative and theoretical analyses. *Journal of Zoology,* 248, 31-48.

Fitch, W.T. 2000. The evolution of speech: A comparative review. *Trends in Cognitive Sciences,* 4, 258-267.

Fitch, W.T., and Giedd, J. 1999. Morphology and development of the human vocal-tract. A study using magnetic resonance imaging. *Journal of the Acoustic Society of America*, 106, 1511-1522.

Fitch, W.T., and Hauser, M.D. 2003. Unpacking ‘honesty’: Vertebrate vocal production and the evolution of acoustic signals. In: *Acoustic Communication.* (Ed. by A.M. Simmons, R.R. Fay, and A.N. Popper), pp. 65-137. New York: Springer.

Folstad, I., and Karter, A.J. 1992. Parasites, bright males, and the immunocompetence handicap. *American Naturalist,* 139, 603-622.

Galeotti, P., Saino, N., Sacchi, R., and Møller, A.P. 1997. Song correlates with social context, testosterone and body condition in male barn swallows. *Animal Behaviour,* 53, 687-700.

Ghazanfar, A.A., and Rendall, D. 2008. Evolution of human vocal production. *Current Biology,* 18, 457-460.

González, J. 2003. Estimation of speaker's weight and height from speech: A reanalysis of data from multiple studies by Lass and colleagues. *Perceptual and Motor Skills,* 96, 297-309.

González, J. 2004. Formant frequencies and body size of a speaker: A weak relationship in adults humans. *Journal of Phonetics,* 32, 277-287.

González, J. 2006. Research in acoustics of human speech sounds: Correlates and perception of speaker body size. *Recent Research Developments in Applied Physics,* 9, 1-15.

González, J., and Oliver, J.C. 2004. Percepción a través de la voz de las características físicas del hablante: identificación de la estatura a partir de una frase o una vocal. *Revista de Psicología General y Aplicada,* 57, 21-33.

Grassi, M. 2005. Do we hear size or sound? Balls dropped on plates. *Perception & Psychophysics,* 67, 274-284.

Greenberg, D.R., and LaPorte, D.J. 1996. Racial differences in body type preferences of men for women. *International Journal of Eating Disorders,* 19, 275-278.

Greisbach, R. 1999. Estimation of speaker height from formant frequencies. *Forensic Linguistics,* 6, 265-277.

Harris, T.R. Fitch, W.T., Goldstein, L.M., and Fashing, P.J. 2006. Black and white colobus monkey (C*olobus guereza)* roars as a source of both honest and exaggerated information about body mass. *Ethology,* 112, 911-920.

Hillenbrand, J., Getty, L.A., Clark, M.J., and Wheeler, K. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustic Society of America,* 97, Pt. 1, 3099-3111.

Hintze, J.L. 1989. *NCSS Version 5.1*. Kayeswille, Utah.

Hollien, H., Green, R., and Massey, K. 1994. Longitudinal research on adolescent voice changes in males. *Journal of the Acoustic Society of America,* 96, 2646-2654.

Hollien, H., and Jackson, B. 1972. Normative data on the speaking fundamental frequency characteristics of young adult males. *Journal of Phonetics,* 1, 117-120.

Hughes, S.M., Harrison, M.A., and Gallup, G.G., Jr. 2002. The sound of symmetry: Voice as a marker of developmental instability. *Evolution and Human Behaviour,* 173-180.

Hughes, S.M., Dispenza, F., and Gallup, G.G., Jr. 2004. Ratings of voice attractiveness predict sexual behaviour and body configuration. *Evolution and Human Behaviour,* 20, 295-304.

Hughes, S.M., and Gallup, G.G., Jr. 2008. Why are we attracted to certain voices? Voice as an evolved medium for the transmission of psychological and biological information. In: *Emotions in the Human Voice: Vol. II. Clinical Evidence.* (Ed. by K. Izdebski), pp. 231-238. California: Plural Publishing, Inc.

Hughes, S.M., Patizzo, M.J., and Gallup, G.G., Jr. 2008. The sound of symmetry revisited. *Journal of Nonverbal Behaviour,* 32, 93-108.

Ives, D.T., Smith, D.R.R., and Patterson, R.D. 2005. Discrimination of speaker size from syllable phrases. *Journal of the Acoustic Society of America,* 118, 3816-3822.

Jackson, L.A., and McGill, O.D. 1996. Body type preferences and body characteristics associations with attractive and unattractive bodies by African Americans and Anglo Americans. *Sex Roles,* 35, 295-307.

Jenkins, J.S. 1998. The voice of the castrato. *Lancet*, 351, 1877-1880.

Johnston, V.S., Hagel, R., Franklin, M., Fink, B., and Grammer, K. 2001. Male facial attractiveness: evidence for hormone-mediated adaptive design. *Evolution and Human Behaviour*, 22, 251-267.

Jolliffe, I.T. 2002. *Principal Component Analysis.* New York: Springer.

Kewley-Port, D. 1995. Thresholds for formant-frequency discrimination of vowels in consonantal context. *Journal of the Acoustic Society of America,* 97, 3139-3146.

Kewley-Port, D., and Watson, C.S. 1994. Formant frequency discrimination for isolated English vowels. *Journal of the Acoustic Society of America,* 95, 485-496.

Kewley-Port, D., Li, Z., Zheng, Y., and Neel, A.. 1996. Fundamental frequency effects on thresholds for vowel formant discrimination. *Journal of the Acoustic Society of America,* 100, 2462-2470.

Klinge, A., and Klump, G.M. 2009. Frequency difference limens of pure tones and harmonics within complex stimuli in Mongolian gerbils and humans. *Journal of the Acoustic Society of America,* 125, 304-314.

Kreiman, J., Gerratt, B.R., and Gabelman, B. 2002. Jitter, shimmer and noise in pathological voice quality perception (A). *Journal of the Acoustic Society of America,* 112, 2446-2446.

Künzel, H.J. 1989. How well does average fundamental frequency correlate with speaker height and weight. *Phonetica,* 46, 117-125.

Ladefoged, P. 2001. *Vowels and Consonants: An Introduction to the Sounds of Languages.* Oxford: Blackwell.

Lattner, S., Meyer, M.E., and Friederici, A.D. 2005. Voice perception: Sex, pitch, and the right hemisphere. *Human Brain Mapping,* 24, 11-20.

Lieberman, P., and Blumstein, S.E. 1988. *Speech Physiology, Speech Perception, and Acoustic Phonetics.* Cambridge, MA: Cambridge University Press.

Macmillan, N.A., and Creelman, C.D. 2005. *Detection Theory: A User's Guide.* Cambridge, MA: Cambridge University Press.

Majewski, W., Hollien, H., and Zalewski, J. 1972. Speaking fundamental frequency of Polish adult males. *Phonetica,* 25, 119-125.

Mermelstein, P. 1978. Difference limens for formant frequencies of steady-state and consonant-bound vowels. *Journal of the Acoustic Society of America,* 63, 572-580.

Miller, G. 2000. *The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature.* New York: Anchor Books.

Moore, B.C.J. 1973. Frequency difference limens for short-duration tones. *Journal of the Acoustic Society of America,* 54, 610-619.

Müller, J. 1848. *The Physiology of the Senses, Voice and Muscular Motion with the Mental Faculties* (Translation). London: W. Baly, Walton and Maberly.

Munhall, K.G., and Byrne, S.K. 2007. Animal communication: Big talkers and small talk. *Current Biology,* 17, R247-R249.

Notelovitz, M. 2002. Androgen effects on bone and muscle. *Fertility and Sterility,* 77, S34-S41.

Ogden, C.L., Fryar, C.D., Carroll, M.D., and Flegal, K.M. 2004. Mean body weight, height and body mass index, United States 1960-2002. *Advance Data from Vital and Health Statistics,* no. 347, 1-17.

Oguchi, T., and Kikuchi, H. 1997. Voice and interpersonal attraction. *Japanese Psychological Research,* 39, 56-61.

Owren, M.J, Berkowitz, M., and Bachorowski, J.-A. 2007. Listeners judge talker sex more efficiently from male than female vowels. *Perception and Psychophysics,* 69, 930- 941.

Owren, M.J., and Cardillo, G. 2006. The relative roles of vowels and consonants in discriminating talker identity versus word meaning. *Journal of the Acoustic Society of America*, 119, 1727-1739.

Owren, M.J., and Rendall, D. 1997. An affect-conditioning model of nonhuman primate vocal signaling. In: *Perspectives in Ethology, Vol. 12 Communication.* (Ed. by D.W. Owings, M.D. Beecher & N.S. Thompson), pp. 299 – 346. New York: Plenum Press.

Owren, M.J., and Rendall, D. 2001. Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology*, 10, 58-71.

Pawlowski, B., Dunbar, R.I.M., and Lipowicz, A. 2000. Evolutionary fitness: Tall men have more reproductive success. *Nature*, 403, 156.

Penton-Voak, I.S., and Perrett, D.I. 2000. Female preference for male faces changes cyclically: Further evidence. *Evolution and Human Behaviour,* 21, 39-28.

Pound, N., Penton-Voak, I.S., and Surridge, A.K. 2009. Testosterone responses to competition in men are related to facial masculinity. *Proceedings of the Royal Society: Biological Sciences,* 276, 153-159.

Puts, D.A. 2005. Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behaviour*, 26, 388-397.

Puts, D.A. 2010. Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution and Human Behaviour,* 31, 157-175.

Puts, D.A., Gaulin, S.J.C., and Verdolini, K. 2006. Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior,* 27, 283-296.

Puts, D., Barndt, J., Welling, L., Dawood, K., and Burris, R. 2010. *Sexual selection and women's voices.* Poster presented at the meeting of the Human Behaviour and Evolution Society, Eugene, OR.

Puts, D.A., Hodges, C.R., Cardenas, R.A., and Gaulin, S.J.C. 2007. Men's voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behaviour,* 28, 340-344.

Reby, D., and McComb, K. 2003. Anatomical constraints generate honesty: Acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour,* 65, 519-530.

Reimers, S., and Maylor, E.A. 2006. Gender effects on reaction time variability and trial-to-trial performance: Reply to Deary and Der (2005). *Aging, Neuropsychology, and Cognition*, 13, 479-489.

Rendall, D., Kollias, S., Ney, C., and Lloyd, P. 2005. Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *Journal of the Acoustic Society of America*, 117, 944-955.

Rendall, D., Notman, H. and Owren, M. 2009. Asymmetries in the individual distinctiveness and maternal recognition of infant contact calls and distress screams in baboons. *Journal of the Acoustic Society of America*, 125, 1792-1805.

Rendall, D., and Owren, M. 2010. Vocalizations as tools for influencing the affect and behaviour of others. In: *Handbook of mammalian vocalization: An integrative neuroscience approach.* (Ed. by S.M. Brudzynski), pp. 177-186. Oxford: Academic Press.

Rendall, D., Vokey, J.R., and Nemeth, C. 2007. Lifting the curtain on the Wizard of Oz: Biased voice-based impressions of speaker size. *Journal of Experimental Psychology – Human Perception and Performance,* 33, 1208-1219.

Rhodes, G. 2005. The evolutionary psychology of facial beauty. *Annual Review of Psychology*, 199-226.

Rhodes, G., Yoshikawa, S., Clark, A., Lee, K., McKay, R., and Akamatsu, S. 2001. Attractiveness of facial averageness and symmetry in non-Western populations: In search of biologically based standards of beauty, *Perception*, 30, 611-625.

Riding, D., Lonsdale, D., and Brown, B. 2006. The effects of average fundamental frequency and variance of fundamental frequency on male vocal attractiveness to women. *Journal of Nonverbal Behaviour*, 30, 55-61.

Riede, T., and Fitch, W.T. 1999. Vocal-tract length and acoustics of vocalization in the domestic dog (*Canis familiaris*). *The Journal of Experimental Biology*, 202, 2859-2869.

Roney, J.R., Hanson, K.N., Durante, K.M., and Maestripieri, D. 2006. Reading men's faces: Women's mate attractiveness judgments track men's testosterone and interest in infants. *Proceedings of the Royal Society: Biological Sciences*, 273, 2169-2175.

Saxton, T.K., Caryl, P.G., and Roberts, S.C. 2006. Vocal and facial attractiveness judgments of children, adolescents and adults: The ontogeny of mate-choice. *Ethology*, 112, 1179-1185.

Sell, A., Bryant, G., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., Krauss, A., and Gurven, M. 2010. Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society: Biological Sciences*, (Published Online, June 16, 2010, doi:10.1098/rspb.2010.0769).

Singh, D. 1993. Adaptive significance of female physical attractiveness: Role of waist-to-hip Ratio. *Journal of Personality and Social Psychology,* 65, 293-307.

Singh, D. 1994. Body fat distribution and perception of desirable female body shape by young black men and women. *International Journal of Eating Disorders,* 16, 289-294.

Sinnott, J.M., Brown, C.H., and Brown, F.E. 1992. Frequency and intensity discrimination in Mongolian gerbils, African monkeys and humans. *Hearing Research,* 59, 205-212.

Sinnott, J.M., Owren, M.J., and Peterson, M.R. 1987. Auditory frequency discrimination in primates: Species differences (*Cercpithecus, Macaca, Homo*). *Journal of Comparative Psychology,* 101, 126-131.

Smith, D.R.R., Patterson, R.D., Turner, R., Kawahara, H. and Irino, T. 2005. The processing and perception of size information in speech sounds. *Journal of the Acoustic Society of America,* 117, 305-318.

Stanislaw, H., and Todorow, N. 1999. Calculation of signal detection theory measures. *Behaviour Research Methods, Instruments, and Computers*, 31, 137-149.

Stevens, K.N. 1998. *Acoustic phonetics.* Cambridge, MA: MIT Press.

Swami, V., and Tovee, M.J. 2009. Big beautiful women: The body size preferences of male fat admirers. *Journal of Sex Research*, 46, 89-96.

Tabachnick, B.G., and Fidell, L.S. 1989. *Using Multivariate Statistics.* New York: Allyn and Bacon.

Terry, L.L. and Vasey, P.L. 2009. Feederism in a woman. *Archives of Sexual Behavior.* (Published Online, December 30, 2009, doi:10.1007/s10508-009-9580-9).

Thompson, E.H., Pleck, J.H., and Ferrera, D.L. 1992. Men and masculinities: Scales for masculinity ideology and masculinity-related constructs. *Sex Roles,* 27, 573-604.

Thornhill, R., and Gangestad, S.W. 1994. Human fluctuating asymmetry and sexual behaviour. *Psychological Science,* 5, 297-302.

Titze, I.R. 1989. Physiological and acoustic differences between male and female voices. *Journal of the Acoustic Society of America,* 85, 1699-1707.

van Dommelen, W.A., and Moxness, B.H. 1995. Acoustic parameters in speaker height and weight identification: Sex-specific behaviour. *Language and Speech,* 38, 267-287.

Varosanec-Skaric, G. 1999. Relation between voice pleasantness and distribution of the spectral energy. In: *Proceedings of the XIVs International Congress of Phonetic Sciences*, San Francisco.

Vukovic, J., Feinberg, D.R., Jones, B.C., DeBruine, L.M., Welling, L.L.M., Little, A.C., and Smith, F.G. 2008. Self-rated attractiveness predicts individual differences in women's preferences for masculine men's voices. *Personality and Individual Differences,* 45, 451-456.

Wier, C.C., Jesteadt, W. and Green, D.M. 1977. Frequency discrimination as a function of frequency and sensation level. *Journal of the Acoustic Society of America,* 61, 178-184.

Wu, K., and Childers, D.G. 1991. Gender recognition from speech. Part I: Coarse analysis. *Journal of the Acoustic Society of America*, 90, 1828-1840.

Yates, A., Edman, J., and Aruguete, M. 2004. Ethnic differences in BMI and body-self-dissatisfaction among Whites, Asian subgroups, Pacific Islanders, and African-Americans. *Journal of Adolescent Health,* 34, 300-307.

Zahavi, A. 1975. Mate selection – a selection for a handicap. *Journal of Theoretical Biology,* 53, 205-214.

Zuckerman, M., and Driver, R.E. 1989. What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behaviour,* 13, 67-82.