Men’s voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men

David Andrew Puts a,⁎, Carolyn R. Hodges b, Rodrigo A. Cárdenas c, Steven J.C. Gaulin b

aNeuroscience Program, Michigan State University, East Lansing, MI 48824, USA
bDepartment of Anthropology, Center for Evolutionary Psychology, University of California, Santa Barbara, Santa Barbara, CA 93117, USA
cDepartment of Psychology, Michigan State University, East Lansing, MI 48824, USA

Initial receipt 15 November 2006; final revision received 4 May 2007

Abstract

Men’s vocal folds and vocal tracts are longer than those of women, resulting in lower fundamental frequency (F0) and closer spacing of formant frequencies (formant dispersion, Df) in men than in women. The evolutionary reasons for these sex differences are uncertain, but some evidence implicates male dominance competition. Previous manipulations of F0 and Df affected perceptions of dominance among men. However, because these acoustic dimensions were manipulated simultaneously, their relative contributions are unclear. In unscripted recordings of men speaking to a competitor, we manipulated F0 and Df independently and by similar perceptual amounts to examine effects on social and physical dominance ratings. Recordings lowered in either F0 or Df were perceived as being produced by more dominant men than were the respective raised recordings. Df had a greater effect than did F0, and both Df and F0 tended to affect physical dominance more than social dominance, although this difference was significant only for Df.

1. Introduction

Large sex differences characterize the human voice and its anatomical substrates. Men’s fundamental frequency (F0), the primary acoustic correlate of pitch, is approximately half as high as women’s (Titze, 2000). F0 is inversely related to vocal fold length, which is 60% longer in men (Titze, 2000), far greater than the roughly 7% difference in height (Gaulin & Boster, 1985). Formant frequencies, which affect perceptions of voice timbre (Cleveland, 1977), are more closely spaced in men (Rendall, Kollias, Ney, & Lloyd, 2005). This spacing, called formant dispersion (Df), is inversely related to vocal tract length (Fitch & Hauser, 1995), which is greater in men, both absolutely and relative to height (Fitch & Giedd, 1999).

Vocal sex differences emerge at puberty when high testosterone levels cause males’ vocal folds and tracts to grow faster than overall body growth (Fitch & Giedd, 1999; Lee, Potamianos, & Narayanan, 1999). No comparable changes occur in females. The evolutionary reasons for these sex differences are unclear, but several lines of evidence indicate that sexual selection specifically shaped men’s voices (Collins, 2000; Feinberg et al., 2006; Feinberg, Jones, Little, Burt, & Perrett, 2005; Puts, 2005; Puts, Gaulin, & Verdolini, 2006).

Sexual selection can operate through mate choice. Women prefer deeper male voices (Collins, 2000; Feinberg et al., 2005; Puts, 2005), especially near ovulation (Feinberg et al., 2006; Puts, 2005) and when evaluating short-term mates (Puts, 2005). Male–male dominance competition may also have played a role (Puts, Gaulin, & Verdolini, 2006). Manipulations of both F0 and Df independently affect women’s ratings of speakers’ age, masculinity, and size (Feinberg et al., 2005), and these parameters might also affect perceptions of dominance among men. Indeed, men rate male voices lowered in both F0 and Df as more dominant than the same voices with these acoustic parameters raised (Puts, Gaulin, & Verdolini, 2006). These manipulations affect ratings of physical dominance (fighting ability) more than ratings of social dominance (respect among peers, leadership, etc.). However, because both F0 and Df were manipulated simultaneously by Puts, Gaulin,
Verdolini, & Hodges, (2006), it is impossible to determine whether \( F_0 \), \( D_0 \), or both, affected perceptions of dominance among men—an issue central to understanding how sexual selection shaped men’s voices. Thus, we examine the relative contributions of \( F_0 \) and \( D_1 \) to ratings of physical and social dominance by manipulating \( F_0 \) and \( D_1 \) independently and by similar perceptual amounts.

2. Methods

2.1. Subjects

Fifty male and 20 female (ages 18–24) English-speaking, normally hearing undergraduates at UC Santa Barbara participated in one of three IRB-approved studies. Two were preliminary just-noticeable-difference (JND) studies; the third was the focal study.

2.2. Studies 1 and 2: JND studies

Experiments were computerized (E-Prime), and voice parameters were manipulated using Praat 4.4.06.

A common measure of perceptual magnitude was required to compare the relative effects of \( F_0 \) and \( D_1 \) on dominance ratings. Using Fechnerian and Weberian psychophysical principles (Marks & Gescheider, 2002), we chose the JND. Although previous studies have examined the JND for various acoustic parameters (e.g., Smith, Patterson, Turner, Kawahara, & Irino, 2005), results vary due to methodological differences. The JND will depend on stimulus magnitude as well as details of instrumentation. Thus, we obtained our own JNDS—tailored to our equipment, methods, and the magnitude of the stimuli used in Study 3—for both \( F_0 \) and \( D_1 \).

Twenty-eight subjects (8 men, 20 women) participated in a JND study of either \( F_0 \) or \( D_1 \). For both \( F_0 \) and \( D_1 \) studies, 30 pairs of voice stimuli were created, comprising an equal mix of 10 specific manipulations, including a null manipulation. \( F_0 \) manipulations ranged from 0.6 to 2.2 semitones in nine increments of 0.2 semitones, plus unmanipulated recordings. \( D_1 \) manipulations ranged from 2.5% to 6.5% changes in apparent vocal tract length in nine increments of 0.5%, plus unmanipulated recordings. For \( F_0 \) manipulation, voices were raised and lowered from baseline levels without affecting tempo. For \( D_1 \) manipulation, formant structure was shifted up and down (increasing or decreasing \( D_1 \)) without affecting tempo. Where \( F_0 \) was affected by this procedure, it was manipulated back, using the procedures described above, to within 0.2 semitones (<0.17 JND) of its original, unmanipulated values. For both \( F_0 \) and \( D_1 \) manipulation, parameters were set to a time step of 0.01, a minimum pitch of 75 Hz, a maximum pitch of 300 Hz, and otherwise to default.

Participants were directed through an experimental session where paired voice stimuli were presented in random order. Each stimulus pair consisted of (A) a 3- to 5-s voice clip of an 18- to 24-year-old man introducing himself, followed by (B) the same clip at the same or different \( F_0 \) or \( D_1 \). AB pairs were presented three times before a response was permitted. Participants were instructed to determine if Clip B was the same as or different from Clip A. We defined the JND as the smallest increment in \( F_0 \) or \( D_1 \) for which 50% of subjects perceived a difference. For \( F_0 \), this was 1.2 semitones; for \( D_1 \), it was a 4% change.

2.3. Study 3: voice effects on dominance

Information on JNDS was used to manipulate 30 voice recordings selected randomly from 111 produced in a previous “dating-game” experiment (Puts, Gaulin, & Verdolini, 2006) where men (ages 18–24) were recorded as they spoke to a male competitor.

In Study 3, \( F_0 \) and \( D_1 \) were both raised and lowered independently using the procedures described above. An experimental manipulation of 1.5 JND was chosen because, from our data, 100% of subjects should perceive differences between voices raised and lowered by this amount in either \( F_0 \) or \( D_1 \). At the same time, this manipulation remains close to the JND, where small increments are maximally likely to have comparable effects on perception (Marks & Gescheider, 2002). For \( F_0 \) manipulation, voices were raised and lowered by 1.8 semitones (1.5 JND) from baseline levels, and for \( D_1 \) manipulation, formant structure was shifted up and down by 6% (1.5 JND), without affecting tempo.

Thus, from each of the 30 original voices, five versions were produced: unmanipulated, raised \( F_0 \), lowered \( F_0 \), raised \( D_1 \), and lowered \( D_1 \), for a total of 150 voice recordings. These recordings were distributed into five stimulus sets of 30 recordings, each set comprising 6 raised \( F_0 \), 6 lowered \( F_0 \), 6 raised \( D_1 \), 6 lowered \( D_1 \), and 6 unmanipulated recordings and only one version of each of the 30 voices. Thus, no subject heard two versions of the same voice.

2.3.1. Procedure

Each male subject \((n=42)\) was seated at a computer station, where a program instructed him on the tasks to be performed, played through headphones the audio stimuli to be rated, and recorded his choices. Subjects rated 30 voices (one stimulus set) on whether each speaker was likely to be able to win physical fights (physical dominance) and whether he was likely to be a respected leader (social dominance). Physical dominance was assessed by selecting from a 10-point scale with endpoints labeled strongly agree and strongly disagree below the statement: “If this man got in a fistfight with an average male undergraduate student, this man would probably win” (Puts, Gaulin, & Verdolini, 2006). Subjects assessed social dominance by selecting from a 10-point scale with endpoints labeled extremely dominant and extremely submissive underneath Mueller and Mazur’s (1997) description: “a dominant person tells other people what to do, is respected, influential, and often a leader; whereas submissive people are not influential or assertive and are usually directed by others.”
To eliminate the influence of any order effects on relationships between voice manipulations and dominance ratings, we presented the 30 stimuli in a different random order for each subject, and whether a given subject always rated social or physical dominance first was determined randomly.

2.3.2. Data treatment

Eight or nine subjects rated each stimulus set, and their ratings were averaged to produce mean physical and social dominance ratings for each recording. Thus, each original voice received 10 mean dominance ratings (5 acoustic variations×2 dominance types). Interrater reliability was high (Cronbach’s \( \alpha = .81 \pm .08 \)). Nevertheless, to control for any global differences in subjects’ rating tendencies, we first standardized each subject’s ratings (converted to \( z \) scores) before they were averaged with other subjects’ ratings. Standardization did not affect the results.

Effects of \( F_0 \) and \( D_f \) on dominance ratings were analyzed using repeated measures ANOVA. All \( p \) values are two tailed; \( \alpha = .05 \).

3. Results

Three factors, each with two levels, were analyzed: manipulation (raised vs. lowered), acoustic measure (\( F_0 \) vs. \( D_f \)), and dominance type (physical vs. social). In separate

<table>
<thead>
<tr>
<th></th>
<th>( D_f )</th>
<th>( F_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Increased</td>
<td>Decreased</td>
</tr>
<tr>
<td>Physical dominance</td>
<td>–.25 (.11)</td>
<td>.19 (.13)</td>
</tr>
<tr>
<td>Social dominance</td>
<td>–.15 (.11)</td>
<td>.10 (.12)</td>
</tr>
</tbody>
</table>

In addition to replicating the finding that vocal masculinity affects dominance perceptions among men (Puts, Gaulin, & Verdolini, 2006), we report three novel results here. First, both \( D_f \) and \( F_0 \) independently affected attributions of dominance among men. Second, to the degree that our ±1.5 JND manipulations of \( D_f \) and \( F_0 \) were perceptually similar, \( D_f \) appears to have a greater effect than \( F_0 \) on dominance ratings. Finally, both acoustic measures tended to more strongly affect attributions of physical dominance than attributions of social dominance, although this difference was significant only for \( D_f \).

These results are likely to have ecological validity. To the extent that men have evolved to assess dominance during
verbal interactions, the stimuli used in this study are realistic (men of the raters’ age speaking spontaneously to a male competitor), and the within-speaker/between-rater design did not draw attention to the acoustic manipulations. Moreover, the effects of within-subject $D_f$ and $F_0$ manipulations were observed in the presence of naturally occurring between-subject variation in speech content and extralinguistic features, such as amplitude and aspiration rate. Thus, $D_f$ and $F_0$ probably affect perceptions of dominance among men in real interactions. This suggests a function for a deep male voice and its anatomical substrates: men’s voices evolved to signal physical dominance.

A potential shortcoming of the present study concerns the distinction between physical and social dominance. Henrich and Gil-White (2001) argue persuasively that social influence can be achieved through force or force threat (what we call here physical dominance). Alternatively, deference may be freely given to individuals who possess valued qualities and who would thus be said to enjoy prestige. Although our variable “social dominance” in some ways approximates prestige (“is respected, influential, and often a leader”), in other respects, it connotes coercion (“tells other people what to do ... assertive”). Thus, depending upon how subjects understood this variable, social dominance may confl ate physical dominance (or simply “dominance”) and prestige. It is possible that $D_f$ affected ratings of respect, influence, and leadership only because this social status variable suggested the threat of force. On the other hand, subjects were also asked to rate physical fighting ability; hence, they may have understood the leadership variable to differ from this. If so, $D_f$ may have affected this variable because physical competitive ability was a prestigious quality to our subjects. Future research should more clearly distinguish between dominance and prestige.

4.1. What does a masculine voice advertise?

Several lines of evidence suggest a relationship between these acoustic parameters and physical competitive ability. First, both $D_f$ and $F_0$ may correlate with body size. Some studies have found statistically significant relationships between $F_0$ and men’s height (Graddol & Swann, 1983) and weight (Evans, Neave, & Wakelin, 2006), although most have not (e.g., Collins, 2000; Kunzel, 1989; Lass & Brown, 1978; Rendall et al., 2005; van Dommelen & Moxness, 1995). Similarly, some studies have found relationships between $D_f$ and men’s height (Evans et al., 2006; Rendall et al., 2005) and weight (Evans et al., 2006; Gonzalez, 2004), although others have not [Collins, 2000 (neither height nor weight); Gonzalez, 2004 (weight but not height); Rendall et al., 2005 (height but not weight)]. Relatively small sample sizes may account for some of these discrepancies. However, if $D_f$ is developmentally constrained to more closely reflect height than is $F_0$ (Fitch & Hauser, 1995), this could explain our observation that $D_f$ more strongly affects dominance attributions.

$F_0$ also appears to correlate negatively with circulating androgens (Dabbs & Mallinger, 1999), which have been positively related to physical aggressiveness (Harris, 1999) and physical prowess (Clark & Henderson, 2003). More generally, it has been suggested that masculine traits such as low $D_f$ and $F_0$, whose development or maintenance depends on high androgen levels, may be honest signals of health and vigor (Folstad & Karter, 1992). Finally, some aspects of men’s voices may reflect self-perceived dominance (Gregory, 1994; Puts, Gaulin, & Verdolini, 2006). For example, men who rate themselves as physically dominant tend to lower their $F_0$ from baseline when competing, whereas men who rate themselves as nondominant tend to raise it (Puts, Gaulin, & Verdolini, 2006).

4.2. Selection for low male voices versus high female voices

$D_f$ and $F_0$ may have evolved to exaggerate the appearance of size in men (Fitch & Giedd, 1999; Morton, 1977). This could occur if there were a strong enough correlation between voice and physical prowess for selection to favor deference to masculine voices. However, current utility is insufficient evidence that a trait is an adaptation; there must also be evidence that the trait was modified for this function (West-Eberhard, 1992). Unfortunately, most vocal structures are soft tissue and would not fossilize, and scant data exist on sex differences in primate vocal anatomy. Thus, evolutionary trends in vocal anatomy cannot easily be established by paleontological or comparative means at present.

On the other hand, ontogeny offers clues. Precipitous pubertal changes suggest that men’s voices have been modified to sound deeper, with developmental events added to elongate the vocal tract and lengthen and thicken the vocal folds relative to overall body size. Of course, it is also possible that such pubertal changes occurred in both sexes ancestrally and that females subsequently lost these developmental events. This could occur, for example, due to selection for neotenic features in women (e.g., Jones, 1995), which might increase their apparent residual reproductive value. Indeed, men may prefer higher female voices (Collins & Missing, 2003). However, sexual selection tends to be stronger among males in mammals generally and in humans in particular (Daly & Wilson, 1988).

4.3. Male contests versus female choice

Men’s voices may have been modified over human evolution to sound deeper, but if so, was the function to increase physical dominance among men or attractiveness to women? In fact, $F_0$ and $D_f$ seem to signal dominance more effectively than they increase attractiveness. Based on a comparison of effect sizes, simultaneous manipulation of $D_f$ and $F_0$ affected men’s judgments of physical dominance nearly 15 times more than they affected fertile–menstrual phase women’s judgments of sexual attractiveness (Puts, Gaulin, & Verdolini, 2006). This does not necessarily imply
that men’s voices affect mating success primarily through dominance. Voice attractiveness could have a larger effect than voice dominance on men’s mating success. However, with age and sociosexuality (attitudes toward uncompromised sex) statistically controlled, physical dominance ratings of a man’s spontaneous speech significantly predicted his number of sex partners over the last year, but sexual attractiveness ratings did not (Puts, Gaulin, Verdolini, & Hodges, 2006).

Although more work is needed, these data suggest that men’s voices may have evolved as dominance signals, and women secondarily evolved preferences for aspects of men’s voices that conveyed information about mate quality. It has been suggested, for example, that men’s voices signal genetic quality (Feinberg et al., 2006; Hughes, Harrison, & Gallup, 2002; Puts, 2005), including perhaps, heritable physical competitive ability. Once evolved, female preferences might then have become complementary selection pressures on men’s voices.

Acknowledgments

We would like to thank Daniel Fessler and two anonymous reviewers for their helpful comments on the manuscript and Drew Rendall for advice on acoustic manipulations.

References


