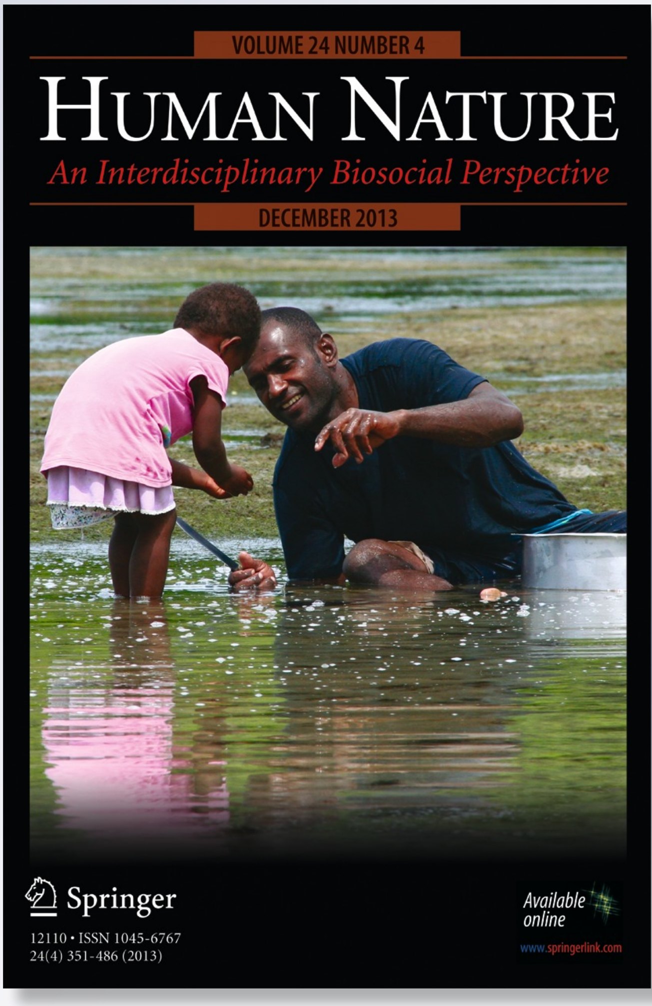
Masculine Men Articulate Less Clearly

Vera Kempe, David A. Puts & Rodrigo A. Cárdenas

Human Nature
An Interdisciplinary Biosocial Perspective

ISSN 1045-6767
Volume 24
Number 4

Hum Nat (2013) 24:461-475
DOI 10.1007/s12110-013-9183-y



Springer

Hum Nat (2013) 24:461–475
DOI 10.1007/s12110-013-9183-y

Masculine Men Articulate Less Clearly

Vera Kempe · David A. Puts · Rodrigo A. Cárdenas

Published online: 15 October 2013

Abstract In previous research, acoustic characteristics of the male voice have been shown to signal various aspects of mate quality and threat potential. But the human voice is also a medium of linguistic communication. The present study explores whether physical and vocal indicators of male mate quality and threat potential are linked to effective communicative behaviors such as vowel differentiation and use of more salient phonetic variants of consonants. We show that physical and vocal indicators of male threat potential, height and formant position, are negatively linked to vowel space size, and that height and levels of circulating testosterone are negatively linked to the use of the aspirated variant of the alveolar stop consonant /t/. Thus, taller, more masculine men display less clarity in their speech and prefer phonetic variants that may be associated with masculine attributes such as toughness. These findings suggest that vocal signals of men's mate quality and/or dominance are not confined to the realm of voice acoustics but extend to other aspects of communicative behavior, even if this means a trade-off with speech patterns that are considered communicatively advantageous, such as clarity and indexical cues to higher social class.

Keywords Vowel space · Allophones · Sex differences

Human voices exhibit sexually dimorphic characteristics such as lower pitch, reduced pitch range, and lower and more closely spaced formants in men (Childers and Wu 1991; Daly and Warren 2001; Puts et al. 2012; Rendall et al. 2005). Recent studies have shown that these characteristics have been subject to sexual selection and tend to signal male prowess, masculinity, and threat potential (Bruckert et al. 2006; Evans et al. 2006; Puts et al. 2012; Wolff and Puts 2010). However, in humans, vocal indicators of prowess, masculinity, and threat potential are displayed in the context of meaningful linguistic communication. It is not clear how these vocal indicators interact with communicatively important features of

V. Kempe (✉)
University of Abertay Dundee, Dundee, UK
e-mail: v.kempe@abertay.ac.uk

D. A. Puts · R. A. Cárdenas
Pennsylvania State University, University Park, PA, USA

speech. Is vocal masculinity associated with increased communicative effectiveness such that linguistic content transmitted by more-masculine-sounding voices is also more easily understood, or do male speakers trade off vocal signals of mate quality and/or dominance with communicatively relevant features of speech? In this study, we explore how physical and vocal indicators of masculinity and threat potential interact with communicative effectiveness and with indexical features of speech that can reveal speaker qualities such as identity, age, social class, or geographical origin.

Vocal Indicators of Masculinity

Lower pitch in males has been associated with several indicators of physical prowess, such as height, weight, and arm strength, and, in some samples, with increased trait-levels of circulating testosterone (Dabbs and Mallinger 1999; Evans et al. 2006; Evans et al. 2008; Puts et al. 2012). Although several studies have not found associations between body size and voice pitch (Collins 2000; Gonzáles 2004; Künzel 1989; Lass and Brown 1978; Rendall et al. 2005; van Dommelen and Moxness 1995), effect sizes are generally similar across studies, and thus a lack of statistical significance may reflect modest effect sizes combined with small sample sizes (range = 15–105). Similarly, male pitch ranges, which tend to be smaller than those of females in normal speech (Daly and Warren 2001) and sometimes even in child-directed speech where exaggerated pitch contours are desirable (Fernald et al. 1989), are linked to self-reported physical aggressiveness (Puts et al. 2012). Studies of the relationship between formant structure and physical attributes of male prowess have yielded controversial results (Fitch 1994, 1997, 2000; Fitch and Giedd 1999; Gonzáles 2004; Rendall et al. 2005), leading some authors to conclude that the relationship between speaker height and vocal tract size is weak (Gonzáles 2004). Another reason for the inconsistent findings is that measures of formant structure such as formant dispersion have failed to account for the relationship between size and the middle formants. To capture the relationship between formant variability and vocal tract size in a manner that accounts for the contribution of all of the first four formants (F1-F4), Puts et al. (2012) introduced the measure of *formant position,* the average formant values standardized using between-sex means and standard deviations, and found that this measure was positively associated with height in both male US college students and male Hadza foragers from Tanzania. Thus, there is mounting evidence that voice acoustics contains cues to male masculinity and threat potential.

Communicative Effectiveness

To explore how vocal signals of masculinity interact with communicative effectiveness, we examine how male speech varies on the continuum from hypo- to hyperspeech (Lindblom 1990): Depending on habitual traits but also on situational and linguistic context, speakers may undershoot or overshoot articulatory targets, which in turn may affect how well they are understood. A famous example relates to the hyperarticulation found in child-directed speech: When mothers address their infants their vowel space is stretched relative to their adult-directed speech, a feature that has been hypothesized to aid infants in distinguishing the various vowel categories of their language, which in turn facilitates language acquisition (Kuhl et al. 1997; Liu et al. 2003). Thus, speakers can adjust articulatory clarity to the needs of the interlocutor to aid in listener understanding. Clarity of speech is determined by a host of contextual and speaker variables, such as the anatomical and motor prerequisites of a speaker but also the context in which utterances are produced. Here we explore how physical and vocal indicators of male mate quality associated with masculinity and threat potential are linked to articulatory clarity. One possibility is that male speakers' articulatory behavior simply corroborates physical and vocal indicators of masculinity and threat potential: The same anatomical constraints that result in more-masculine-sounding voices (i.e., longer vocal tract length) may also affect the size of the vowel space. Such an assertion is supported by the observation that females, who have shorter vocal tracts, also have larger vowel spaces (Fant 1966, 1975; Simpson 2009). If vocal tract anatomy is the main determinant of vowel space in adult-directed speech, then physical and vocal indicators of masculinity should be negatively linked to vowel space size; in other words, more masculine men are expected to produce less differentiated vowels.

On the other hand, speakers may try to offset negative attributions associated with facial and vocal masculinity, such as infidelity (O'Connor et al. 2011) or lack of investment in social relationships and offspring care (Perrett et al. 1998), through learned vocal behavior that signals empathy, agreeableness, and social skill. Such a scenario would lead to the prediction that speakers with more masculine physical and vocal characteristics are more likely to employ hyperarticulation in a compensatory way to counteract the potentially adverse auditory and social effects of longer vocal tracts.

Indexical Features

In addition to transmitting the information contained in the message, human speech also signals indexical information about the speaker, such as gender, age, geographical origin, and social class membership. Social, regional, and age information can be obtained from syntactic constructions such as "I like them bananas" or lexical items such as "outwith" (a Scotticism) or "irregardless" (a word originating in western Indiana), but phonetic information probably provides the most telling indexical cues. For example, in many varieties of English, word-final /t/ can be realized through three allophonic variants: an aspirated voiceless alveolar stop, an unreleased voiceless alveolar stop, and a glottal stop. American English encompasses considerable variability in phonetic variants of /t/, with use of glottal stops being somewhat more pronounced in speakers in their teens and twenties originating from western regions of the US (Eddington and Channer 2010) and aspirated (t) being associated with higher social status (Macaulay 1977; Milroy et al. 1994; Trudgill 1974). The social indexing potential of /t/ is even more pronounced in British English, where the aspirated variant is a feature of Received Pronunciation, an accent associated with deliberate speech clarity and high social prestige. Indeed, in the Tyneside region of England, use of glottal stops instead of a released (t) betrays a speaker's socioeconomic status (Foulkes et al. 2005), a quality that may be relevant in the context of mate choice.

However, speakers can manipulate the indexical features of their speech to accomplish social and interpersonal alignment—for example, by adopting phonetic, syntactic, or lexical features used or valued by their interlocutors (Trudgill 1986:8). In this respect, aspiration in word-final /t/s can serve to signal a speaker's desire for clarity and/or social group affiliation. Thus, if indexical information simply corroborates physical and vocal indicators of masculinity and threat potential, then more-masculine speakers should produce more unreleased and glottalized variants of /t/, which are less clear and may signal younger age and lower-class membership. If, on the other hand, more-masculine men exploit phonetic variability to compensate for the negative attributions associated with masculinity, then they may be more likely to produce an aspirated word-final (t), the perceptually most salient /t/ variant, to signal an effort to be clear and to project higher social class membership and, perhaps, maturity.

To explore whether and how physical and vocal masculinity is linked to phonetic characteristics of speech we examined vowel spaces and /t/ variants in a sample of young male speakers of American English and linked them to physical and vocal indicators of masculinity and threat potential, such as height, weight, arm strength, physical aggressiveness, as well as circulating testosterone. We also included 2D:4D digit ratio, a retrospective marker of prenatal testosterone exposure, because even though the early organizing effects of testosterone do not seem to affect vocal parameters such as pitch or formant dispersion directly (Evans et al. 2008), there is some evidence that 2D:4D may be linked to variation in communicatively relevant vocal features of speech (Kempe 2009). We measured vowel space area and frequency of aspirated (t)s, which can be taken as measures of articulatory clarity and effort in speaking. In addition, production of aspirated (t)s as opposed to unreleased (t)s or glottal stops may serve as an indexical cue to speaker age and social class.

Method

Participants

One-hundred and seventy-six male students from a large northeastern US university participated in this IRB-approved study. The male participants ranged from 18 to 26 years of age with a mean of 20.1 years. Since ethnic, geographical, and psychosexual factors are known to affect socio-phonetic speech profiles, only the 155 participants who identified themselves as white and heterosexual were retained for further analyses. These participants originated predominantly from Midwestern regions of the US.

In addition, 129 normally cycling female students from the same population were tested to obtain between-sex means of formant frequencies required for the calculation of formant position (P_f). The female participants ranged from 18 to 24 years of age with a mean of 19.6 years. To maintain compatibility with the male participants, only the 114 females who identified themselves as white and heterosexual were retained in the sample.

Procedures

Male participants were scheduled for 1-h morning (beginning between 8:20 and 10:00) and evening (beginning between 17:20 and 19:00) sessions 1 week apart. Female participants were scheduled based on self-reported menstrual cycle data to attend sessions between 13:00 and 16:00, one during the late follicular phase and one during the mid-luteal phase. Session order (morning or evening first vs. follicular or mid-luteal first) was randomized. Acoustic, phonetic, anthropometric, hormonal, and psychometric data were collected at both sessions.

Voice Recording

Participants were recorded reading the word list *beat, bit, bet, bait, bat, but, bout, bye, book, boot, boat, bought, bird, car,* and *ago* in an anechoic, soundproof booth using a Shure SM58 vocal cardioid microphone. A curved wire projection from the microphone stand kept the participant's mouth approximately 9.5 cm from the microphone. Voices were recorded into a computer using GOLDWAVE software in mono at a sampling rate of 44,100 Hz and 16 bit quantization, and saved as uncompressed ".wav" files. Recordings were analyzed using Praat software (Boersma and Weenink 2011).

Acoustic Measures

P_f was determined using the method outlined in Puts et al. (2012): We first measured F1 through F4 at each glottal pulse (automated detection by Praat; mean glottal pulses per recording = 552.74±201.11) and averaged across measurements. This method samples a wide range of vocal tract configurations of voiced speech avoiding fricatives. Because Praat occasionally shifts formants (e.g., misattributing F2 as F1), we omitted all formant measurements from glottal pulses for which any value exceeded a predetermined threshold (less than 2% of pulses). Thresholds were based on published data (Rendall et al. 2005) (1,000, 2,850, 3,750, and 4,500 Hz for F1 through F4 for males; 1,250, 3,350, 4,150, and 5,100 Hz for F1 through F4 for females) and were selected to eliminate only clearly erroneous measurements. P_f was computed as the average standardized formant value for the first four formants, where formants were standardized using between-sex means and standard deviations.

So as not to bias the mean or standard deviation toward either sex during standardization, we obtained means and standard deviations using bootstrapping methods, randomly selecting (with replacement) 10,000 samples of 114 men and 114 women. With both sexes included, we obtained mean F1=545.3±74.2 Hz, mean F2=1585.5 ±154.8 Hz, mean F3=2608.7±181.7 Hz, and mean F4=3613.5±271.6 Hz. These values were then used to standardize formants for all participants.

Phonetic Measures

To determine size of the vowel space area, a measure of articulatory clarity and precision (Bradlow et al. 1996), we adopted the procedure outlined in the cross-linguistic study by Kuhl et al. (1997) for vowel spaces encompassed by the vowels /a/, /i/, and /u/: We measured F1 and F2 of the vowels /i:/ in *beat*, /u:/ in *boot*, and /^/ in *but* produced by the male participants. Formants were sampled during a period that was judged by visual inspection to be maximally steady. The vowel /^/ was used as a substitute for /a:/, which was not available in the word list. This resulted in a somewhat compressed vowel space but did not affect comparison of vowel space areas across participants. Misattributions of F1 and F2 affected formant measurements in five cases in session 1 and in three cases in session 2; these cases were excluded from analysis.

Vowel space can be viewed as the area of a triangle encompassed by the three pairs of (F1,F2) coordinates on an *x,y* plane and is calculated using the formula

$$\left[\mathrm{F1}_{/\mathrm{i:}/}{*}\left(\mathrm{F2}_{/\wedge/}-\mathrm{F2}_{/\mathrm{u:}/}\right)+\mathrm{F1}_{/\wedge/}{*}\left(\mathrm{F2}_{/\mathrm{u:}/}-\mathrm{F2}_{/\mathrm{i:}/}\right)+\mathrm{F1}_{/\mathrm{u:}/}{*}\left(\mathrm{F2}_{/\mathrm{i:}/}-\mathrm{F2}_{/\wedge/}\right)\right]\Big/2$$

where $\mathrm{F1}_{/\mathrm{i:}/}$ is the F1 value of the vowel /i:/, $\mathrm{F1}_{/\wedge/}$ is the F1 value for the vowel /^/, and so on.[1]

To determine presence or absence of the aspirated variant of (t), a voiceless stop, we inspected the wave form of the words *beat, bit, bet, bait, bat, but, bout, boot, boat,* and *bought.* The criterion for presence of the aspirated voiceless stop (t) was evidence for a release burst followed by a period of turbulence indicative of aspiration. Absence of this pattern in the signal indicated the presence of an unreleased (t) or a glottal stop.

Anthropometric Measurements

Flexed biceps circumference was measured at its widest point for left and right arms using a tape measure. Left and right hand strengths were obtained using a JAMAR hydraulic hand dynamometer. Biceps size and hand strength are good predictors of overall upper body strength (Sell et al. 2010). Height was measured from a meter stick affixed to a wall, and weight was obtained using an electronic scale. Length of the left and right index (D2) and ring (D4) fingers was measured from photocopies of the palms of participants' hands. Because finger lengths were measured from tip to basal crease, participants' basal creases were marked with a fine-tipped marker to increase visibility. Two trained researchers measured each photocopy with digital callipers (precise to 0.01 mm) for a total of four measurements (2 sessions × 2 measurers) of each finger. If the standard deviation of the four measurements exceeded 1.2 mm, the measurements were examined for errors, which, if detected, were corrected. For both 2D and 4D, all correlations between measurements across raters and across sessions were $r>0.97$. 2D:4D digit ratio was calculated by dividing D2 length by D4 length for each hand.

Testosterone Assays

Saliva was collected for testosterone (T) assays during morning and evening sessions. Contamination of saliva samples was minimized by having participants not eat, drink

[1] We did not use a normalization procedure because vowel-intrinsic normalization, such as MEL-transformation, has been shown to distort speaker-specific anatomical/physiological variation in the acoustic representations of vowels (Adank et al. 2004). However, repeating the analyses with MEL-transformed F1 and F2 values yielded nearly identical results.

(except plain water), smoke, chew gum, or brush their teeth for 1 h before their session. Participants rinsed their mouths with water before chewing a piece of sugar-free Trident gum (inert in salivary hormone assays) to stimulate saliva flow. Approximately 9 ml of saliva was collected in a sodium-azide-coated polystyrene tube. The tube was capped and left upright at room temperature for 18–24 h to allow mucins to settle. Tubes were then frozen at −20 °C until hormone analysis.

We obtained salivary unbound ("free") T concentrations, which correlate strongly with serum T concentrations (Wang et al. 1981). The Salivary Radioimmunoassay Laboratory at the University of Western Ontario performed T radioimmunoassay on 330 male saliva samples, 175 from session 1 and 155 from session 2. All samples went through double ether extraction, followed by radioimmunoassay in duplicate using a Coat-A-Count kit for total T (Diagnostic Products, Los Angeles, CA, USA), modified for use with saliva (for details, see Moffat and Hampson 1996). The average intra-assay coefficient of variation was 6.3%, and sensitivity was 5–10 pg ml^{-1}.

Psychometric Testing

Following anthropometry and saliva collection, each participant completed the Buss and Perry (1992) Aggression Questionnaire, whose 29 items include nine targeting physical aggression (e.g., "Once in a while I can't control the urge to strike another person"). Items are assessed on a 5-point scale anchored at "extremely uncharacteristic of me" and "extremely characteristic of me." Scores for the nine items targeting physical aggression were summed for a composite score of physical aggression.

Results

An analysis of the association between women's body size, hormonal status, and vocal parameters was beyond the scope of this paper and will be reported elsewhere. For the men, correlations of anthropometric measurements across sessions ranged from r=0.667 for left arm biceps circumference to r=0.997 for height. Correlations between left and right side of the body ranged from r=0.736 for 2D:4D to r=0.877 for biceps circumference. Correlations for acoustic parameters across sessions ranged from r=0.511 for F_0-s.d. to r=0.858 for F3. Correlations for phonetic measures across sessions were r=0.542 for vowel space areas and r=0.810 for frequency of (t). All correlations reported here were significant at the p<0.001 level, justifying using averages of all measurements across sides of the body, where appropriate, and across sessions.

T concentrations for morning and evening sessions were modestly but significantly correlated (r=0.530), reflecting temporal variation in T secretion. A comparison of T concentrations from the morning and evening sessions showed that this variation followed the expected diurnal pattern with higher levels in the morning (paired t_{137}=9.93, p<0.0001).

Correlations Between Measurements

F_0-s.d., weight, physical aggression, and T levels were log-transformed to correct for positive skew. Hand strength and biceps circumference were standardized and averaged to produce the composite measure "arm strength." Table 1 shows the zero-order correlations between all measurements. Not surprisingly, height and weight were positively correlated with arm strength. There was also a positive correlation between F_0 and F_0-s.d. Moreover, height, weight, and arm strength were negatively correlated with formant position, suggesting that a more masculine formant position was associated with increased physical threat potential. Finally, we found a negative correlation between circulating T levels and F_0, indicating that higher T levels were associated with deeper voices.

Predictors of Male Vocal Parameters

To explore the independent effects of physical and psychological threat-potential indicators on vocal parameters, and to reduce the risk of Type I error, we entered height, weight, arm strength, physical aggression score, T, and 2D:4D as predictors in a series of multiple regression analyses, with each of the three vocal acoustic parameters (F_0, F_0-s.d., and P_f) as criterion variables, controlling for the effects of the remaining two vocal acoustic parameters. This analysis constitutes an attempt to replicate the findings reported in Puts et al. (2012) for vocal parameters obtained from roughly the same sample of men but from a different speech episode. One difference from the previous study lies in the addition of 2D:4D as a predictor variable to clarify organizational effects of prenatal T because associations between this variable and vocal parameters had been inconsistent in previous research (Evans et al. 2008). The results of the regression analyses are shown in Table 2. All variance inflation factors (VIFs) were below 1.8.

The regression analyses broadly replicate the findings reported in Puts et al. (2012). As shown by the zero-order correlations, height was negatively associated with F_0 and P_f, confirming that taller men tend to have lower pitch and a more masculine distribution of formants, and circulating T was negatively associated with F_0. Arm strength was negatively linked with P_f and not, as in Puts et al. (2012), with F_0-s.d., and the physical aggression score did not have any significant effect. Despite

Table 1 Zero-order correlations between anthropometric, acoustic, psychometric, and phonetic variables and testosterone levels

		2	3	4	5	6	7	8	9
1	Height	0.38**	0.18*	0.14	−0.16†	−0.01	−0.31**	0.07	−0.08
2	Weight	–	0.59**	0.08	0.04	0.05	−0.19*	0.16*	−0.03
3	Arm strength		–	−0.11	−0.06	−0.04	−0.26**	0.30**	0.04
4	2D:4D			–	−0.06	0.08	−0.08	−0.10	−0.04
5	F_0				–	0.37**	0.12	−0.03	−0.21*
6	F_0-s.d.					–	0.09	−0.13	0.04
7	P_f						–	−0.12	0.07
8	Physical aggression							–	−0.06
9	T level								–

† $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 2 Standardized regression coefficients and *t* values in multiple regression analyses with indicators of masculinity and threat potential as predictors of men's vocal acoustic parameters

	F_0		F_0-s.d.		P_f	
	t	β	*t*	β	*t*	β
Height	−2.47*	−0.20	0.75	0.07	−3.13**	−0.27
Weight	1.77†	0.17	0.17	0.02	0.69	0.07
Arm strength	−1.10	−0.11	0.04	<0.01	−2.49*	−0.25
Physical aggression	0.10	<0.01	−1.06	−0.09	−0.51	−0.04
2D:4D	−1.25	−0.09	1.17	0.09	−1.06	−0.08
Testosterone	−3.29**	−0.24	1.59	0.13	0.71	0.06
F_0			4.97***	0.40	0.45	0.04
F_0-s.d.	4.97***	0.37			0.47	0.04
P_f	0.45	0.04	0.47	0.04		
Overall adjusted R^2	0.200***		0.124**		0.112**	

† $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

these small deviations from the previous findings, this new data set generally confirms that such indicators of masculinity and threat potential as body size, strength, and circulating T are linked to acoustic features of male voices.

Predictors of Vowel Space

Size of area encompassed by the F1-F2 coordinates of the vowels/u:/,/i:/, and/^/was entered as criterion variable into a multiple regression with the anthropometric, acoustic, psychometric, and hormonal measurements as predictor variables. The results, given in Table 3, showed that P_f significantly predicted vowel space, indicating that men with a more masculine formant position also had smaller vowel spaces. In addition, we found a negative relationship between height and vowel space such that taller men had less-differentiated vowels. Note that while height predicted P_f, as demonstrated in Table 2, the multiple regression results demonstrate that height also had an independent relationship with vowel space over and above the relationship between P_f and vowel space. It is important to remember that P_f was computed based on all voiced segments of speech whereas vowel space was computed only based on the steady-state portions of the vowels /i:/, /u:/, and /^/. This suggests that in addition to displaying a more masculine position of formants on the male–female continuum in general, presumably owing to their longer vocal tracts, taller men also produce less clearly differentiated vowels (Fig. 1).

Predictors of Allophonic Variation in (t)

Frequency of use of aspirated (t) was bimodally distributed, with 40 men never producing an aspirated (t), 39 men producing an aspirated (t) in all 20 words across both sessions, and 76 men producing an aspirated (t) in some, but not all, instances. To examine the link between the anthropometric, acoustic, psychometric, and hormonal predictor variables

Table 3 Standardized regression coefficients and Wald χ^2 values in a multinomial regression with frequency of aspirated (t) as criterion variable, and standardized regression coefficients and *t* values in a linear multiple regression with vowel space as criterion variable, with anthropometric, acoustic, and hormonal variables as predictors in both models

	Frequency of aspirated (t)		Vowel space	
	Wald χ^2	β	*t*	β
Height	3.85*	−0.46	−2.39*	−0.21
Weight	0.08	0.06	0.86	0.09
Arm strength	1.02	0.19	1.59	0.16
Physical aggressiveness	1.68	−0.20	−1.42	−0.11
2D:4D	1.11	0.16	0.97	0.08
Testosterone	4.72*	−0.32	0.45	0.03
F_0	0.03	−0.11	−0.19	−0.02
F_0-s.d.	0.31	0.14	0.06	0.01
P_f	0.38	−0.30	4.63***	0.37

* $p<0.05$, *** $p<0.001$

and the frequency of productions of aspirated (t), we used a multinomial regression with a logit link function. The results, also provided in Table 3, indicate that height and circulating T were negatively associated with lower frequency of use of aspirated (t). Thus, taller, more masculine men were less likely to produce aspirated (t)s.

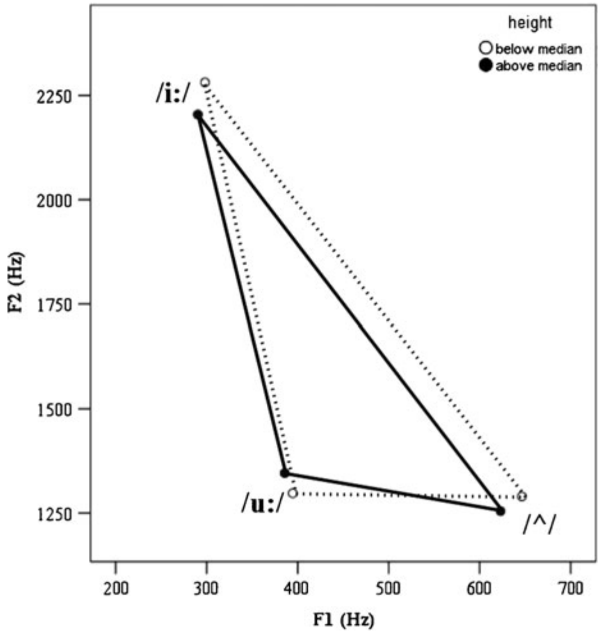


Fig. 1 Vowel spaces of men above and below the median height

One question that arises is whether allophonic choice was determined by temporal T fluctuations or by trait-level T. To examine this question we took advantage of the observed diurnal variation in circulating T: If T fluctuations affect choice of aspirated (t) we should observe a lower frequency of aspirated (t) during the high-T morning sessions compared to the low-T evening sessions. A paired-samples *t*-test comparing frequency of aspirated (t) between morning and evening sessions revealed that this was not the case ($p=0.8$). This suggests that it is trait-level T, an indicator of masculinity, that affects the extent to which male speakers produce aspirated (t)s.

Discussion

Our results confirm previous findings of links between physical indicators of masculinity and features of the male voice (Bruckert et al. 2006; Evans et al. 2006; Puts et al. 2012; Wolff and Puts 2010), although some small discrepancies with respect to the specific links between physical and vocal features remain. For example, we did not observe the negative correlation between circulating testosterone and the composite measure of formant frequency and formant dispersion reported elsewhere (Bruckert et al. 2006). Crucially, however, our results go beyond these previous findings by showing that physical and vocal indicators of masculinity and threat potential are linked to communicatively relevant features of speech such as vowel space and choice of allophonic variants. Specifically, we found that taller men produced less differentiated vowels and were also less likely to produce the perceptually clearer aspirated variant of /t/. Moreover, more masculine formant position was associated with smaller vowel space, and higher circulating testosterone was associated with more frequent use of glottalized and unreleased /t/s. These findings suggest that physical and vocal indicators of masculinity are associated with speech patterns that are not only lower in clarity but also may project younger age and lower social class membership. We note that no association was found between 2D:4D digit ratio, a marker of prenatal testosterone, and any of the vocal parameters, suggesting that organizational effects of testosterone appear to affect neither vocal indicators of masculinity nor phonetic variability in speech.

Thus, our findings did not support the idea that more masculine men use communicatively relevant speech patterns to offset some of the potentially negative impressions associated with masculinity, such as infidelity threat or low paternal investment (O'Connor et al. 2011; Perrett et al. 1998). Indeed, there may not be sufficient pressure for men to compensate for such potentially negative impressions by using clearer and more socially desirable speech patterns given that vocal masculinity is perceived as indicative of other desirable attributes, such as leadership quality (Klofstad et al. 2012; Puts et al. 2006, 2007; Tigue et al. 2012). On the contrary, the phonetic patterns of speech seem to corroborate the physical and vocal indicators of masculinity, at least in speech produced in a neutral laboratory setting. It appears, then, that men readily trade in phonetic attributes of clear and cooperative communication for displays of vocal masculinity that, in turn, signal such qualities as threat potential and leadership. These findings can be interpreted in the context of the distinction between dominance (social status achieved through the ability to use force) and prestige (status freely conferred by other group members) (Henrich and Gil-White 2001): If vocal indicators of masculinity are signals of dominance, and allophonic variants associated with maturity and higher social class affiliation are indicators of prestige, then our results suggest that young men may tend to resolve trade-offs between signaling dominance and signaling prestige in favor of signaling dominance.

To what extent are the observed links between physical and vocal indicators of masculinity and phonetic speech patterns a by-product of anatomical features such as greater height and vocal tract length? While this study is to our knowledge the first one to demonstrate systematic within-sex differences in vowel space, there is considerable evidence for between-sex differences in vowel space, with women exhibiting larger vowel spaces than men (Fant 1966, 1975). This difference arises because upward shifts in F1 and F2, the main determinants of vowel quality, are not uniform in females but affect each vowel differently. There is debate as to whether this non-uniform shift is attributable to biomechanical constraints or to learned behaviors. On the one hand, differences in vocal tract size are assumed to lead to differences in articulatory dimensions, to differences in outcomes of articulatory gestures when performed at same speed as in males, as well as to different harmonic consequences of the interaction between pitch and articulation (Simpson 2001, 2009). On the other hand, the evidence for sociophonetic variability in speech suggests that speech patterns are learned behaviors and can be deployed depending on a host of factors, such as age, sex, sexual orientation, hormonal status, and social class. For example, homosexual American men and women display non-uniform changes in vowel space, which differ from those of heterosexual members of their own and the opposite sex (Pierrehumbert et al. 2004). Furthermore, prepubescent girls produce more exaggerated differences in voice onset time (VOT) between voiced and voiceless plosives than boys and younger or older girls (Whiteside and Marshall 2001; Whiteside et al. 2004b). In addition, VOTs have been shown to vary during the menstrual cycle, with VOT differences between voiced and voiceless plosives being more pronounced when estrogen and progesterone levels are highest (Wadnerkar et al. 2006; Whiteside et al. 2004a). Thus, female speech clarity is enhanced during the luteal phase of the menstrual cycle, an observation that cannot be explained by cyclical changes in peripheral tissues resulting in stiffness of the vocal folds (Simpson 2009).

For the present findings, we cannot rule out that changes in vowel space as a function of height are the by-product of biomechanical factors. However, the finding that height and testosterone levels predicted reduced production of aspirated /t/s points to a behavioral explanation of the observed speech patterns, as it is difficult to envisage biological factors that may govern the choice of these allophones. A number of reasons may explain why more masculine men show a tendency to articulate less clearly: Perhaps reduced speech clarity is a consequence of reduced effort expended on communication if masculine men perceive themselves to be dominant or more desirable mates. It is also possible that the choice of glottal stops may serve to signal membership in social groups that value masculine attributes such as toughness and lack of empathy or cooperation. Whatever the underlying proximate mechanism, the present study shows that masculinity and threat potential are signaled not just by features of the voice but also by patterns of speech.

One open question is whether these systematic, within-sex differences in vowel space and allophonic variation do indeed have an effect on speech intelligibility and communicative effectiveness. It is well known that listeners compensate for variability in pitch and timbral features of speech through speaker normalization (Ladefoged and Broadbent 1957). Despite normalization, however, speakers still differ in perceived intelligibility. Kwon (2010) demonstrated that female speakers display superior speech intelligibility, which was linked to, among other features, increased vowel space. Extrapolating this finding to within-sex differences in vowel space leads to the prediction, to be tested in future speech perception studies, that shorter men with larger vowel spaces should produce more intelligible speech.

In addition to affecting speech intelligibility and leaking masculinity and threat potential, allophonic variation may also be evaluated by females differentially depending on mating context. For example, small vowel spaces may be treated by female listeners as potential signals of low paternal investment: Because variability in speaker vowel space could affect the acquisition of speech sounds in infants (Liu et al. 2003), men with smaller vowel spaces may be perceived to be less-investing, and hence less-desirable, fathers; after all, their speech input to prospective offspring may be of somewhat lower quality. Consequently, vowel space size should be more likely to affect voice attractiveness judgments adversely in long-term, as opposed to short-term, mating contexts (Puts 2005), a hypothesis that will have to be tested in future research.

References

Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America, 116*, 3099–3107. doi:10.1121/1.1795335.

Boersma, P., Weenink, D. (2011). Praat: Doing phonetics by computer [Computer program] Version 5316, retrieved May 30, 2012 from http://www.fon.hum.uva.nl/praat/.

Bradlow, A., Torretta, G., & Pisoni, D. (1996). Intelligibility of normal speech I: global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20*, 255–272. doi:10.1016/S0167-6393(96)00063-5.

Bruckert, L., Lienard, J. S., Lacroix, A., Kreutzer, M., & Leboucher, G. (2006). Women use voice parameters to assess men's characteristics. *Proceedings of the Royal Society B-Biological Sciences, 273*, 83–89. doi:10.1098/rspb.2005.3265.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63*, 452–459. doi:10.1037//0022-3514.63.3.452.

Childers, D. G., & Wu, K. (1991). Gender recognition from speech. Part II: fine analysis. *Journal of the Acoustical Society of America, 90*, 1841–1856. doi:10.1121/1.401664.

Collins, S. (2000). Men's voices and women's choices. *Animal Behaviour, 60*, 773–780. doi:10.1006/anbe.2000.1523.

Dabbs, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences, 27*, 801–804. doi:10.1016/S0191-8869(98)00272-4.

Daly, N., & Warren, P. (2001). Pitching it differently in New Zealand English: speaker sex and intonation patters. *Journal of Sociolinguistics, 5*, 85–96. doi:10.1111/1467-9481.00139.

Eddington, D., & Channer, C. (2010). American English has goʔ a loʔ of glottal stops: social diffusion and linguistic motivation. *American Speech, 85*, 338–351. doi:10.1215/00031283-2010-019.

Evans, S., Neave, N., & Wakelin, D. (2006). Relationships between vocal characteristics and body size and shape in human males: an evolutionary explanation for a deep male voice. *Biological Psychology, 72*, 160–163. doi:10.1016/j.biopsycho.2005.09.003.

Evans, S., Neave, N., Wakelin, D., & Hamilton. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology and Behavior, 93*, 783–788. doi:10.1016/j.physbeh.2007.11.033.

Fant, G. (1966). A note on vocal tract size factors and non-uniform F-pattern scalings. *STL- Quarterly Progress and Status Report, 4*, 22–30.

Fant, G. (1975). Non-uniform vowel normalization. *STL-Quarterly Progress and Status Report, 2*, 1–19.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language, 16*, 477–501. doi:10.1017/S0305000900010679.

Fitch, T. W. (1994). *Vocal tract length perception and the evolution of language.* Unpublished Ph.D. thesis, Brown University.

Fitch, T. W. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America, 102*, 1213–1222. doi:10.1121/1.421048.

Fitch, T. W. (2000). Evolution of speech: a comparative review. *Trends in Cognitive Science, 4*, 258–267. doi:10.1016/S1364-6613(00)01494-7.

Fitch, T. W., & Giedd, J. (1999). Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *Journal of the Acoustical Society of America, 106*, 1511–1522. doi:10.1121/1.427148.

Foulkes, P., Docherty, G., & Watt, D. (2005). Phonological variation in child-directed speech. *Language, 81*, 177–206. doi:10.1353/lan.2005.0018.

Gonzáles, J. (2004). Formant frequencies and body size of speaker: a weak relationship in adult humans. *Journal of Phonetics, 32*, 277–287. doi:10.1016/S0095-4470(03)00049-4.

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 21*, 165–196. doi:10.1016/S1090-5138(00)00071-4.

Kempe, V. (2009). Child-directed speech prosody in adolescents: relationship to 2D:4D, empathy, and attitudes towards children. *Personality and Individual Differences, 47*, 610–615.

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B-Biological Sciences, 279*, 2698–2704. doi:10.1098/rspb.2012.0311.

Kuhl, P. K., Andruski, J. E., Christovich, I. A., Christovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*, 684–686. doi:10.1126/science.277.5326.684.

Künzel, H. J. (1989). How well does average fundamental frequency correlate with speaker height and weight? *Phonetica, 46*, 117–125. doi:10.1159/000261832.

Kwon, H. (2010). Gender difference in speech intelligibility using speech intelligibility tests and acoustic analyses. *Journal of Advanced Prostodontics, 2*, 71–76. doi:10.4047/jap.2010.2.3.71.

Ladefoged, P., & Broadbent, D. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America, 29*, 98–104. doi:10.1121/1.1908694.

Lass, N. J., & Brown, W. S. (1978). Correlational study of speakers' heights, weights, body surface areas, and speaking fundamental frequencies. *Journal of the Acoustical Society of America, 63*, 1218–1220. doi:10.1121/1.381808.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H and H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Amsterdam: Kluwer. doi:10.1007/978-94-009-2037-8_16.

Liu, H.-M., Kuhl, P. K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infant speech discrimination skills. *Developmental Science, 6*, F1–F10. doi:10.1111/1467-7687.00275.

Macaulay, R. K. S. (1977). *Language, social class, and education: A Glasgow study.* Edinburgh: The University Press.

Milroy, J., Milroy, L., Hartley, S., & Walshaw, D. (1994). Glottal stops and tyneside glottalization: competing patterns of variation and change in British English. *Language Variation and Change, 6*, 327–357. doi:10.1017/S095439450000171X.

Moffat, S. D., & Hampson, E. (1996). Salivary testosterone levels in left- and right-handed adults. *Neuropsychologia, 34*, 225–233. doi:10.1016/0028-3932(95)00090-9.

O'Connor, J. J. M., Re, D. E., & Feinberg, D. R. (2011). Voice pitch influences perceptions of sexual infidelity. *Evolutionary Psychology, 9*, 64–78.

Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*, 884–887. doi:10.1038/29772.

Pierrehumbert, J. B., Bent, T., Munson, B., Bradlow, A. R., & Bailey, M. J. (2004). The influence of sexual orientation on vowel production (L). *Journal of the Acoustical Society of America, 116*, 1905–1908. doi:10.1121/1.1788729.

Puts, D. A. (2005). Mating context and menstrual phase affect female preferences for male voice pitch. *Evolution and Human Behavior, 26*, 388–397. doi:10.1016/j.evolhumbehav.2005.03.001.

Puts, D. A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior, 27*, 283–296. doi:10.1016/j.evolhumbehav.2005.11.003.

Puts, D. A., Hodges, C., Cárdenas, R. A., & Gaulin, S. J. C. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior, 28*, 340–344. doi:10.1016/j.evolhumbehav.2007.05.002.

Puts, D. A., Apicella, C. L., & Cárdenas, R. A. (2012). Masculine voices signal men's threat potential in forager and industrial societies. *Proceedings of the Royal Society B-Biological Sciences, 279*, 601–609. doi:10.1098/rspb.2011.0829.

Rendall, D., Kollias, S., Ney, C., & 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 Acoustical Society of America, 117*, 944–955. doi:10.1121/1.1848011.

Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., Krauss, A., & Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society B-Biological Sciences, 277*, 3509–3518. doi:10.1098/rspb.2010.0769.

Simpson, A. P. (2001). Dynamic consequences of differences in male and female vocal tract dimensions. *Journal of the Acoustical Society of America, 109*, 2153–2164. doi:10.1121/1.1356020.

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass, 3*, 621–640. doi:10.1111/j.1749-818X.2009.00125.x.

Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behaviour. *Evolution and Human Behavior, 33*, 210–216. doi:10.1016/j.evolhumbehav.2011.09.004.

Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, P. (1986). *Dialects in contact*. Oxford: Blackwell.

Van Dommelen, W. A., & Moxness, B. H. (1995). Acoustic parameters in speaker height and weight identification: sex-specific behaviour. *Language and Speech, 38*, 267–287.

Wadnerkar, M. B., Cowell, P. E., & Whiteside, S. P. (2006). Speech across the menstrual cycle: a replication and extension study. *Neuroscience Letters, 408*, 21–24. doi:10.1016/j.neulet.2006.07.032.

Wang, C., Plymate, S., Nieschlag, E., & Paulsen, C. A. (1981). Salivary testosterone in men: further evidence of a direct correlation with free serum testosterone. *Journal of Clinical Endocrinology and Metabolism, 53*, 1021–1024. doi:10.1210/jcem-53-5-1021.

Whiteside, S. P., & Marshall, J. (2001). Developmental trends in Voice Onset Time: some evidence for sex differences. *Phonetica, 58*, 196–210. doi:10.1159/000056199.

Whiteside, S. P., Hanson, A., & Cowell, P. E. (2004a). Hormones and temporal components of speech: sex differences and effects of menstrual cyclicity on speech. *Neuroscience Letters, 367*, 44–47. doi:10.1016/j.neulet.2004.05.076.

Whiteside, S. P., Henry, L., & Dobbin, R. (2004b). Sex differences in voice onset time: a developmental study of phonetic context effects in British English. *Journal of the Acoustical Society of America, 116*, 1179–1183. doi:10.1121/1.1768256.

Wolff, S. E., & Puts, D. A. (2010). Vocal masculinity is a robust dominance signal in men. *Behavioral Ecology and Sociobiology, 64*, 1673–1683. doi:10.1007/s00265-010-0981-5.

Vera Kempe is professor of psychology of language learning at the University of Abertay Dundee. She is interested in evolutionary constraints on language learning and processing.

David Puts is an associate professor of anthropology at Pennsylvania State University. He is interested in sexual selection and the processes of sexual differentiation.

Rodrigo Cárdenas is an NSF postdoctoral fellow in the Department of Psychology at Pennsylvania State University. He is interested in the evolution of human cognition.