ANIMAL BEHAVIOUR, 2000, **60**, 773–780
doi:10.1006/anbe.2000.1523, available online at http://www.idealibrary.com on IDEAL®

Men's voices and women's choices

SARAH A. COLLINS
Behavioural Biology Section, Institute of Evolutionary & Ecological Sciences, Leiden, The Netherlands

*(Received 21 March 2000; initial acceptance 10 May 2000;
final acceptance 5 July 2000; MS. number: 6530)*

I investigated the relationship between male human vocal characteristics and female judgements about the speaker. Thirty-four males were recorded uttering five vowels and measures were taken, from power spectrums, of the first five harmonic frequencies, overall peak frequency and formant frequencies (emphasized, resonance, frequencies within the vowel). Male body measures were also taken (age, weight, height, and hip and shoulder width) and the men were asked whether they had chest hair. The recordings were then played to female judges, who were asked to rate the males' attractiveness, age, weight and height, and to estimate the muscularity of the speaker and whether he had a hairy chest. Men with voices in which there were closely spaced, low-frequency harmonics were judged as being more attractive, older and heavier, more likely to have a hairy chest and of a more muscular body type. There was no relationship between any vocal and body characteristic. The judges' estimates were incorrect except for weight. They showed extremely strong agreement on all judgements. The results imply that there could be sexual selection through female choice for male vocal characteristics, deeper voices being preferred. However, the function of the preference is unclear given that the estimates were generally incorrect.

Relationships between acoustic parameters and body size (Appleby & Redpath 1997; Giacoma et al. 1997), hormonal status (Fusani et al. 1994; Beani et al. 1995) and age (Ballintijn & ten Cate 1997) have been found in a number of taxa (birds, anurans and some mammals). However, the relationships between acoustic parameters and individual characteristics in primates (including humans) appear to be more complex (Lass & Brown 1978; van Dommelen 1993; Hauser 1993).

One commonly found relationship, in nonprimate species, is between the acoustic frequency of a vocalization and body size: larger individuals produce lower-frequency sounds (Morton 1977; Howard & Young 1998). The relationship between body size and frequency holds across primate species (Hauser 1993), but not within species (Fitch 1997). Most people believe that a deep human voice (i.e. a low-frequency voice) indicates a larger person, although this is not true (van Dommelen & Moxness 1995). As pointed out by Fitch (1997), this is not surprising when one considers the structure of the vocal apparatus in primates. The fundamental frequency of the voice, in both nonhuman primates and humans, is dependent upon the thickness and size of the vocal folds (Fant 1960; Lieberman 1984; Schön Ybarra 1995) which are soft tissues. Testosterone increases the size and thickness of the folds (Hollien 1960; Beckford et al. 1985). Changes in the vocal apparatus occur at the same time as changes in body shape and size during puberty, but the relative increase in the size of the vocal folds is greater than that in body size (Hollien et al. 1994) and is independent of measures of skeletal size (Beckford et al. 1985). During adolescence, there is a correlation between vocal frequency and body size, because body size increases and the voice deepens simultaneously, although the correlation disappears during the latter stages of puberty (Hollien et al. 1994).

However, some information about body size is contained in the vocalizations of primates. The vocal tract is made up of hard tissue, its length being related to both skull and skeletal size (Fitch & Hauser 1995; Fitch 1997) and the size of the tract determines the resonance frequencies of calls (Fant 1960; Lieberman & Blumstein 1988; Moore 1992). The resonance frequencies, known as formant frequencies, are emphasized frequencies within vocalizations (Fig. 1). Fitch (1997) showed that in rhesus macaques, *Macaca mulatta*, the length of the vocal tract and the formant frequencies produced are both related to body size. The same is probably true in baboons (Owren et al. 1997). More precisely, it is the difference in frequency between successive formant frequencies (formant dispersion) that correlates with body size. Larger individuals have smaller differences between the formant frequencies (Fitch 1997). This information could be used by other individuals to infer body size.

Correspondence and present address: S. A. Collins, Behavioural Ecology Research Group, School of Life and Environmental Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, U.K. (email: Sarah.Collins@nottingham.ac.uk).

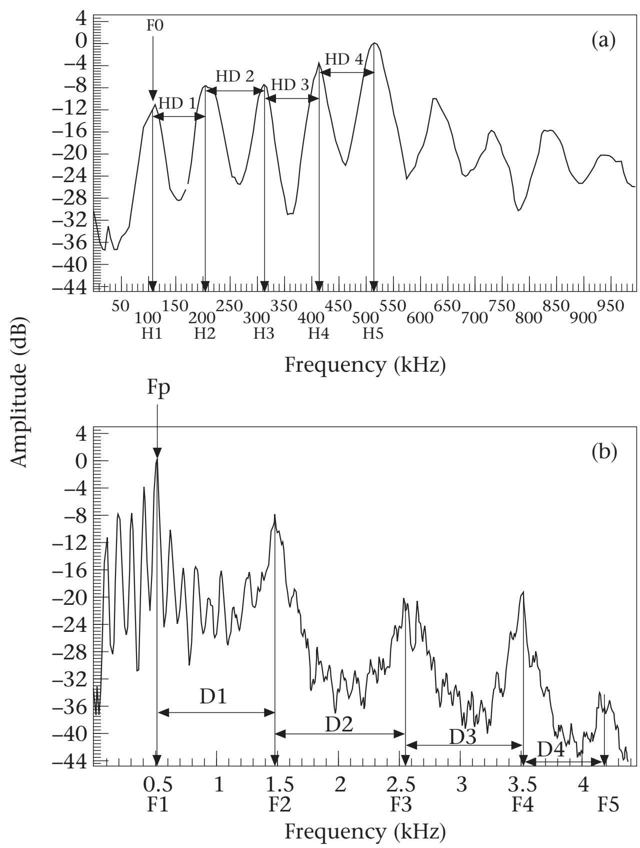

Figure 1. Power spectrums of the vowel O. (a) HD n=harmonic difference (harmonic spacing is the mean of HD 1–5); Hn=harmonic peak frequency; F0=fundamental frequency. (b) Fp=peak frequency; Fn=formant frequency; Dn=difference in formants (formant dispersion is the mean of D 1–5).

Vowels are produced with an open vocal tract (Baer et al. 1991), so the acoustic parameters are determined less by the position of the tongue and soft palate, and more by the length and shape of the vocal tract (Baer et al. 1991; Moore 1992; Maurer et al. 1996). Since vowels contain formant frequencies, information on body size may be conveyed by human vowel sounds.

A great deal of work has been done on how characteristics of the human voice influence judgements about speakers. It is known that sex (Bennett & Montero-Diaz 1982; Childers & Wu 1991, from vowels alone), age (Linville 1996; Mulac & Giles 1996), and race (Lass et al. 1979; Walton & Orlikoff 1994) can be identified by listeners. The cues that are used to assess the sex of the speaker are related to both the fundamental (lowest frequency produced; Fig. 1) and formant frequencies (Whiteside 1998a, b). Men have lower formant frequencies that are closer together (Childers & Wu 1991), that is, there is less formant dispersion (Fitch 1997). Listeners also make judgements about the relative masculinity of male speech. Speech judged less masculine has higher second formant frequencies of some vowels, and differences in the frequency change across vowels (Avery & Liss 1996).

van Dommelen & Moxness (1995) investigated whether listeners could judge the height and weight of speakers, from recitations of words or short paragraphs (see also Lass & Davies 1976; Lass et al. 1980). Male height and weight were estimated accurately, especially by men. However, these estimates were based on the rate of speech, although listeners 'thought' they were basing their judgement on the frequency of the voice (van Dommelen & Moxness 1995). None of the frequency parameters was related to weight or height.

No work has been conducted on human female preference for male voices. However, there has been extensive investigation of the effect of male characteristics, other than voice, on female mate choice (review Barber 1995). My aims were (1) to investigate whether male body size can be assessed by vocal characteristics such as formant dispersion and (2) to evaluate agreement between female judgements about individual characteristics based on listening to male voices, a prerequisite for sexual selection through female choice of male voice.

METHODS

Subjects and Recordings

The speakers were 34 males aged 18–30 years (mean 22.41 years), all Dutch natives. Voices were recorded by three groups of experimenters using a Sony TC D5 tape recorder and Sennheiser MKH 70 microphone. The microphone was held at 30 cm and a constant recording sound level was used. Each male was asked to say the five vowels five times in their normal voice, at a constant speed: A, as in normal; E, as in say; I, as in see; O, as in no; U, duel (the Dutch U is not a sound used in English; duel is the closest approximation).

Groups 1 and 2 consisted of 10 men and group 3 of 14 men. When the subjects were recorded, a number of body measures were taken: weight (range 51–89 kg), shoulder width (34–45 cm), hip width (31–43 cm), chest circumference (82–103 cm) and height (1.64–2.05 m). The hip to shoulder width ratio (0.97–1.41) was also calculated. Group 1 males were asked whether they had chest hair (6/10 men answered 'yes').

Analysis of Vowels

Logarithmic Power Spectrums were made of the vowels (A E I O U=one series) using the Avisoft SAS lab sound analysis program (sampling frequency 11 kHz, smoothed over 20 Hz for clarity, and normalized to maximum). For analysis, I used only the vowel series 2, 3 and 4, as speakers tended to intonate series 1 and 5 differently.

The following measures were calculated from the spectrum (Fig. 1): peak frequency of each the first five harmonics (harmonic peaks), overall peak frequency (overall peak), mean spacing in frequency between the harmonics (harmonic spacing) and the peak frequency of the first three, four or five formant frequencies (in some subjects the higher formant frequencies were not measurable for certain vowels). For the first two parameters I calculated the average over all five vowels in each of the three sets. For harmonic spacing and formant dispersion a mean was taken for each vowel in each series, and then an overall mean across the three series. For individuals with few, or no, fourth and fifth formant frequencies, I calculated formant dispersion with all available data for an individual (as in Fitch 1997). The average formant frequencies (1–5) over the three sets of vowels were calculated for each individual letter.

In addition, I calculated the coefficient of variation (variance/mean × 100) for the harmonic spacing within each individual, as I hypothesized that what is often described as voice roughness (variable harmonics) may be used as a judgement criterion.

In speech analysis the Linear Predictive Coding technique is generally used to measure formant frequencies (e.g. Childers & Wu 1991; Fitch 1997). However, the mean formant frequencies calculated here are similar to those found using this technique.

Table 1. Results of ANCOVA (correcting for group membership) or Pearson correlation analysis, all groups combined, N=34 all tests

Variables	R value (unless stated otherwise)	P
Weight and chest	0.75	<0.001
Weight and hip	r^2=0.71	<0.001
Chest and hip	r^2=0.25	0.002
Weight and height	0.51	0.002
Peak frequency and all harmonics	0.534–0.594	0.001
Harmonic spacing and harmonics (1st–5th)	0.985–0.998	<0.001
Harmonics 1 to 5 with each other	0.987–0.997	<0.001
Harmonic spacing and peak frequency	0.59	<0.001

Stimulus Series

One stimulus tape was made from each of the three subject groups. A series of vowels (one of the middle three series) from each speaker was recorded on to the 'test' tape, the speed of utterance of the vowels and the amplitude equalized in Avisoft. For each group of judges, the subject voices were recorded on to the preference tape in five different speaker orders. The vowels themselves were always in the same order as initially recorded, A E I O U.

Preference Tests

As judges, I used 54 native Dutch women, aged 18–30 (mean age 21.1 years), none of whom knew the speakers. There were three groups of judges: group 1 (22 judges) and group 2 (21 judges), each of which heard a preference tape with 10 voices, and group 3 (11 judges) who listened to the tape with 14 voices. For groups 1 and 2, the preference tests were conducted by two experimenters in two different rooms, so in effect there was a group 1a (10 judges) and 1b (12 judges), and group 2a (10 judges) and 2b (11 judges).

All tests were conducted in a quiet room; the stimulus voices were played twice through headphones at constant amplitude. Each order of voices was used in rotation within a group. Judges were given a questionnaire and asked to score the following: (1) the attractiveness of the man to whom the voice belonged (scale of 1–10: 10 being the most attractive), (2) the speaker's weight and (3) his age (they were told the men were between 18 and 30). Some groups were asked additional questions: group 3 were asked to guess the body type (scale of 1–3; from slim to muscled); group 1 (a and b) were asked to guess whether the speaker had a hairy chest (yes=1, no=0) and to estimate the speaker's height (1a only).

Analysis of Judgements

The estimate for body type was calculated as the mean judgement of the 11 judges (range 1.2–2.4). A chest hair estimate score was calculated from the total number of 'yes' judgements for each voice in group 1 (range 1–18). For each voice the mean attractiveness, estimated weight, height and age were calculated. However, for weight, height and age, subjects began with an estimate for the first voice heard, then increased or decreased for subsequent voices. Thus, the range in estimations for different women within a group varied, for example from 70 to 80 kg for one judge and from 80 to 100 kg for another. Therefore, a Kendall rank score (from the Kendall coefficient of concordance) for all the judges' estimates except chest hair and body type, was calculated. This ranks a judge's scores from lowest to highest and takes the mean rank for each speaker from all judges. The scores from group 3 were corrected (score × 10/14), as there were 14 voices and, therefore, 14 ranks rather than 10.

Statistical Analysis

Some male body measures differed between the three groups (one-way ANOVA), chest and hip width, and age. Therefore I used ANCOVA (controlling for group) to look for relationships within body measures containing these parameters, and Pearson correlations otherwise. There were strong correlations within both body and vocal characters (Table 1). Spearman correlation was used for chest and body type estimates.

Two principal component analyses with varimax rotation were performed to reduce the number of variables, first on the body measures (not including age):

Component 1=body size; explains 44% of the variance; heavy, tall men with large hips and shoulders have a high score.

Component 2=body shape; 30.8% of the variance; men with large shoulders and small hips have a high score.

Second, on the vocal parameters:

Component 1=harmonic peaks and overall peak frequency; explains 29% of the variance; voices with more widely spaced harmonics and higher peak frequencies have a high score.

Component 2=third and 'I' formant frequencies; explains 15.2% of the variance; voices with higher formant frequencies have a high score.

Component 3=first and 'A' formant frequencies; explains 13.5% of the variance; voices with higher formant frequencies have a high score.

Component 4='O' formant frequencies and formant dispersion; explains 13.3% of the variance; voices with higher formant frequencies and wider formant dispersion have a high score.

The fourth and fifth formant frequencies were not in the principal component analysis as they were not present for all speakers. One speaker (group 3) had missing third formant frequencies on two letters, so no components were calculated for this male. Therefore, in analyses involving one of the four components there is one fewer data point, that is, 13 or 33 rather than 14 or 34.

Agreement between judges was calculated with the Kendall coefficient of concordance, except for chest hair estimates (as scores are 1 or 0). For this parameter the number of 'yes' answers were added up for each of the two groups of 10 women who heard group 1 voices, and agreement between the two groups analysed by Spearman correlations. All analyses were conducted with SPSS version 8.0. All significant results are reported as such after Bonferroni corrections. Tests are two tailed.

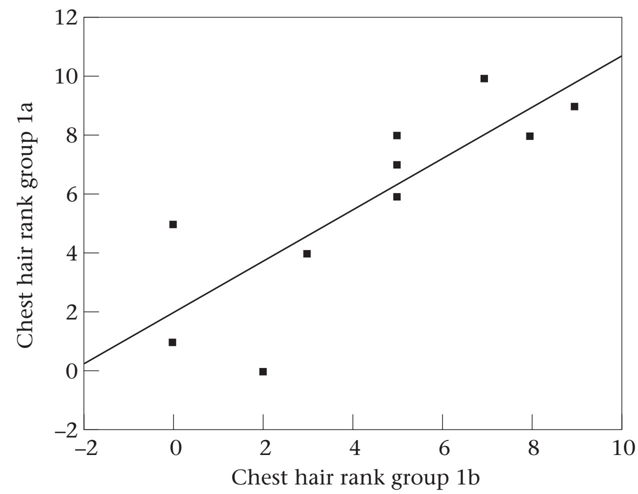

Figure 2. The relationship between groups 1a and 1b in the number of judges estimating whether a particular man had chest hair or not (N=10).

Table 2. Kendall coefficients of concordance, a measure of the agreement between judges

	Attractiveness	Weight	Age
Group 1			
χ^2	25.70	65.90	66.30
W	0.238	0.610	0.614
P	<0.005	<0.001	<0.001
Group 2			
χ^2	58.12	61.57	57.24
W	0.307	0.342	0.303
P	<0.001	<0.001	<0.001
Group 3			
χ^2	38.70	47.19	32.87
W	0.298	0.363	0.253
P	<0.001	<0.001	<0.005

N=10 subjects in groups 1 and 2 and 14 in group 3.

RESULTS

Acoustic Parameters and Body Characteristics

Regression analyses were conducted with the body components as independent variables and the vocal components and voice roughness (coefficient of variance in harmonics) as dependent variables. In addition, regressions on to the body components were conducted with formant dispersion calculated from all formant frequencies, and from only formant frequencies 3–5 (less likely to be determined by the specific vowel in question). None of the regression analyses was significant.

Acoustic Parameters and Listener Judgements

Agreement between judges

The judges in groups 1a and 1b agreed on whether a voice belonged to a man with chest hair (Spearman correlation coefficient: r_s=0.864, N=10, P<0.001; Fig. 2). In all groups judges agreed on the attractiveness, weight and age of the speaker (Kendall coefficient of concordance: Table 2). Judges (group 2) also agreed on the body type (W=0.19, N=14, P<0.05) and height of the speaker group 1a: W=0.23, N=10, P<0.05). For subsequent analyses, the subgroups (a, b) are combined within groups 1 and 2. In all three groups together the judgments of attractiveness, age and weight (Kendall rank values) were significantly correlated with each other (age–weight, r=0.73, N=34, P<0.001: age–attractiveness, r=0.46, N=34, P<0.01: weight–attractiveness, r=0.56, N=34, P<0.01).

Judgements and acoustic characteristics

All regression analyses on judgements were conducted with the four vocal characteristic components as independent variables and the Kendall rank judgement scores as dependent variables. The harmonics factor was related negatively to a number of different judgements. Voices with smaller differences between the harmonics, and lower frequency harmonics (the harmonics component), were judged as: more attractive (regression: $F_{1,32}$=7.5, P<0.01, R^2=0.17; Fig. 3a); heavier ($F_{1,32}$=27.02, P<0.001, R^2=0.45; Fig. 3b); older ($F_{1,32}$=15.9, P<0.01, R^2=0.32; Fig. 3c); more likely to have chest hair (r_s= −0.855, N=10, P<0.003; Fig. 3d), group 1; taller ($F_{1,8}$=6.9, P<0.05, R^2=0.40), group 1a; and more muscular (r_s= −0.69, N=13, P<0.01), group 3.

Men judged as more attractive may be estimated as heavier, taller and older because of their presumed attractiveness. However, in a stepwise regression the attractiveness rank was not significantly related to estimated weight or age when the harmonics component was entered.

Judgements and male physical characteristics

The estimated weight was predicted by the actual weight of the speaker ($F_{1,32}$=10.4, P=0.005, R^2=0.22;

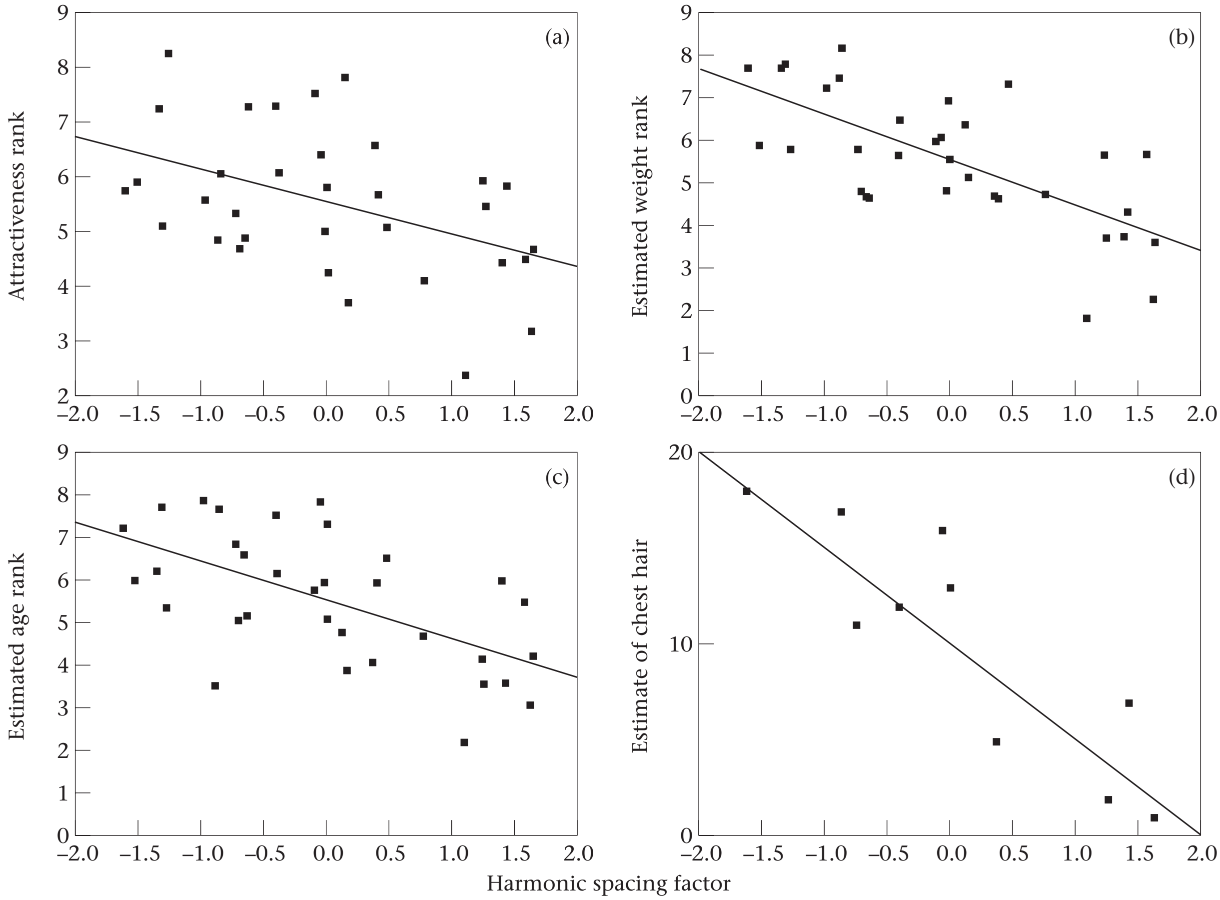

Figure 3. The relationship between the harmonics component and (a) attractiveness rank; (b) estimated weight rank; (c) estimated age rank; and (d) estimated presence of chest hair. (a–c) N=34, (d) N=10. (Low scores on the component mean close together harmonics.)

Fig. 4), even though there was no relationship between the harmonics component (related to estimated weight) and actual body weight ($F_{1,31}$=0.84, NS). The estimated age of the speaker was not predicted by the age of the speaker, in fact in two of the three groups the relationship was negative. The number of judges guessing a speaker had chest hair was not different between men with and without chest hair (Mann–Whitney U test: U=8, N_1=4, N_2=6, NS). There was no relationship between estimated and the actual height (group 1a; Spearman correlation: r_s=0.06, N=10, NS).

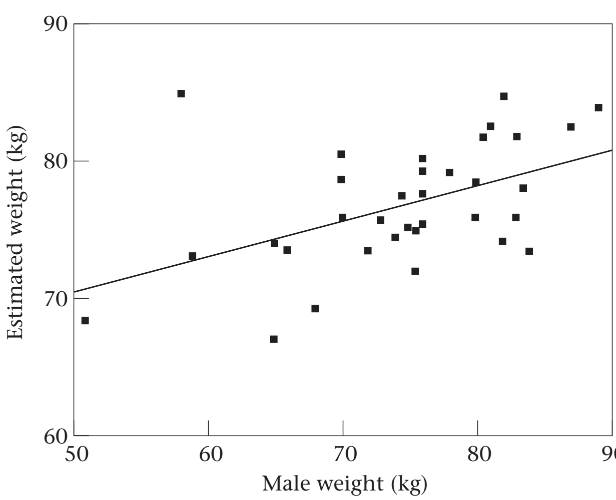

Figure 4. The relationship between estimated weight and actual weight of subject; N=34.

DISCUSSION

There are four main results. First there was no correlation between male vocal and body characteristics. Second, women made a number of judgements about men based on their spoken vowel sounds. In general, women found men's voices with harmonics that are closer together and lower in frequency more attractive. They also estimated these men as being heavier, older, more likely to have a hairy chest and more likely to have a muscular body type. Third, women strongly agreed about these physical characteristics associated with a male voice. Finally, weight was reliably estimated from the voice, but age, height and hairiness of chest were not. Given the results of previous studies, it is unlikely that these results apply only to Dutch women judges and Dutch male speakers.

I shall first discuss the relationship (or lack of) between body characteristics and vocal characteristics of the speaker. As stated earlier, less masculine sounding speakers have higher formant frequencies, among other characteristics (Avery & Liss 1996), but there was no relationship between formant frequencies and body characteristics or judgements in this study. By far the best predictor of the formant frequencies in humans is the particular vowel being spoken (Maurer et al. 1996). However, formant dispersion might still indicate body size, as the length of the vocal tract should constrain formant production. A possible explanation for the lack of relationship was offered by Fitch (1994): in humans the vocal tract may be disassociated from skeletal size because the larynx has descended to a position deeper in the throat. This descent is relatively greater in men than women (Fitch 1994). However, it seems that although men have deeper voices with lower formant dispersion, and heavier bodies than women, within men there is no relationship between a larger body and a voice with lower formant dispersion.

I hypothesized that there may be a relationship between indicators of testosterone level and fundamental (lowest) frequency. While I did not measure testosterone directly, subjects were asked whether they had chest hair, and their hip and shoulder widths were measured and used to calculate a shoulder:hip ratio, which is related to levels of testosterone (Kasperk et al. 1997). However, I found no association between the above and any measures of vocal characteristics. Perhaps measures of actual testosterone levels, in particular during adolescence, may be related to the fundamental frequency of the voice, through its effect on vocal fold measures (Hollien 1960; Beckford et al. 1985).

The judgements made by the women about the voices were highly dependent upon male vocal characteristics. Although there was no relationship between body and vocal characteristics, women used vocal characteristics to infer physical characters. Voices with close harmonic structure and lower frequency harmonics were perceived as belonging to older, heavier, taller, hairier and more attractive men.

The estimated and actual weights were correlated suggesting that the women were using an honest cue of body size in the male voice to estimate weight. However, the strongest predictor of estimated weight was the harmonics component, which was not correlated with body size. Perhaps this component provides a good approximation to the weight of the speaker, or there might have been vocal cues not measured that were related to speaker weight. For all other judgements, the women in this study were incorrect about the actual characteristics of the speaker. Simply raising the vocal frequency is not enough to render male speech less masculine sounding (Childers & Wu 1991), further evidence that voice 'depth' does not allow assessment of masculinity.

So why do women make these incorrect judgments and why use harmonic characteristics as a guide? One possible explanation is a kind of peak shift effect (Weary et al. 1993). Peak shift causes a stronger response to stimuli that are at the extremes of the distributions of each category, furthest from the mean for the alternative category. It is possible that male voices with parameters towards the boundaries of the distribution for men are furthest from the mean parameters for women. In effect, they are seen as being more masculine. They may also be assumed to have body characteristics that are more masculine. A second explanation may simply be that the categorical differences between men and women might be taken to imply correlations within the categories, such that men with deeper voices are assumed to be more masculine, a cognitive rather than perceptual effect (Wittenbrink et al. 1998).

Third, Brunswick (1955) suggested that perception of objects (sounds in this case) is based on a complex series of cues, many of which are unreliable. Since cues are probabilistic (not fully dependable), a perceiver must make a perceptual compromise between the cues in order to make a judgment. The perceiver must still arrive at rapid and generally valid judgements regarding the object. In this experiment women used the depth of vocal frequency to judge weight (high functional validity), although this was not ecologically valid (to use Brunswick's terminology).

Another possibility is that cultural traditions give rise to the perceived association. In movies and television larger men have deeper voices. This may reflect preconceptions, or may give rise to those preconceptions. Whichever is cause and which effect, it seems certain that larger men are expected to have deeper voices. In a recent film magazine (Park 1999) it was stated that for the voices of both Darth Vader and the current Star Wars villain, a 'voice over' was used, the reason being that although the actors had large bodies their voices were not deep enough to be convincing. This kind of reasoning could give rise to cultural expectations of an association between body size and a low-frequency voice.

A further explanation could be that there was an association between the harmonic frequency of vocalizations and body characteristics before the development of language. These now unreliable judgements might have been useful in evolutionary history.

It is hard to imagine a situation where women cannot visually as well as aurally interact with a man, but this does not mean that vocal cues are not used in mate choice. Scent is important in mate choice (Wedekind & Furi 1997) because it provides information not available from visual cues. The same may be true for vocal cues.

Since nonhuman primate vocalizations, such as the grunts of baboons and coos of macaques, are similar to prelinguistic human vocalizations (Locke & Snow 1997), the same components probably influenced the evolution of the frequency characteristics. In nonhuman primates, vocalizations are important in male–male competition, and may have evolved to communicate over long distances (Hauser 1993; Mitani & Stuht 1998). In previous studies men were better at estimating male weight from vocalizations than women, perhaps because of the importance of accurately assessing opponents during intrasexual competition (Fitch & Hauser 1995). However, it is in the interests of females to pay attention to vocalizations as an indicator of the dominance rank of a male (for references see Qvarnstrom & Forsgren 1998). The deepening of human male voices compared to women may be due to selection pressure from competition with other ‘males, the environment, and a weaker selection pressure from women's preference for deeper voices. In future work we shall look at judgements of male voices by males, to see if the same judgements are made.

Acknowledgments

We thank all the students who enthusiastically collected the data, Marc Steigenga, Sebastian Ruinard, Henk den Bakker, Sarah Bollendorff, Marjolein van Ruiyven, Helen Moed, Gideon Bevelander, Erik Antonissen, Natascha Wallaard, Maaike Bruinsma, Guus Zijlstra, Karin Weening, Alison Bartlett and Judith Betenburg and also Hans Slabbekoorn and two anonymous referees who helped with comments on the analysis. The work was supported by Leiden University.

References

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

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

Baer, T., Gore, J. C., Gracco, L. C. & Nye, P. W. 1991. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels. *Journal of the Acoustical Society of America*, **90**, 799–828.

Ballintijn, M. R. & ten Cate, C. 1997. Sex differences in the vocalizations and syrinx of the collared dove (*Streptopelia decaocto*). *Auk*, **114**, 22–39.

Barber, N. 1995. The evolutionary psychology of physical attractiveness: sexual selection and human morphology. *Ethology and Sociobiology*, **16**, 395–424.

Beani, L., Panzica, G., Persichella, P. & Dessi Fulghieri, F. 1995. Testosterone induced changes of call structure, midbrain and syrinx anatomy in partridges. *Physiology and Behavior*, **58**, 1149–1157.

Beckford, N. S., Schain, D., Roor, S. R. & Schanbacher, B. 1985. Androgen stimulation and laryngeal development. *Annals of Otology, Rhinology and Laryngology*, **94**, 634–640.

Bennett, S. & Montero-Diaz, L. 1982. Children's perception of speaker sex. *Journal of Phonetics*, **10**, 113–121.

Brunswick, E. 1955. Representative design and probabilistic theory in a functional psychology. *Psychology Review*, **62**, 193–217.

Brunswick, E. 1956. *Perception and the Representative Design of Psychological experiments*. Berkeley, California: University of California Press.

Childers, D. G. & Wu, K. 1991. Gender recognition from speech 2. Fine analysis. *Journal of the Acoustical Society of America*, **90**, 1841–1856.

van Dommelen, W. A. 1993. Speaker height and weight identification: a reevaluation of some old data. *Journal of Phonetics*, **21**, 337–341.

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.

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

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

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

Fitch, W. T. & Hauser, M. D. 1995. Vocal production in nonhuman-primates: acoustics, physiology and functional constraints on honest advertisement. *American Journal of Primatology*, **37**, 191–219.

Fusani, L., Beani, L. & Dessi Fulghieri, F. 1994. Testosterone affects the acoustic structure of the male call in the grey partridge (*Perdix perdix*). *Behaviour*, **128**, 301–310.

Giacoma, C., Zugolaro, C. & Beani, L. 1997. The advertisement calls of the green toad (*Bufo viridis*): variability and role in mate choice. *Herpetologican*, **53**, 454–464.

Hauser, M. D. 1993. The evolution of nonhuman primate vocalisations: effects of phylogeny, body weight and social-context. *American Naturalist*, **142**, 528–542.

Hollien, H. 1960. Some laryngeal correlates of vocal pitch. *Journal of Speech and Hearing Research*, **3**, 52–58.

Hollien, H., Green, R. & Massey, K. 1994. Longitudinal research on adolescent voice change in males. *Journal of the Acoustical Society of America*, **96**, 2646–2654.

Howard, R. D. & Young, J. R. 1998. Individual variation in male vocal traits and female mating preferences in *Bufo americanus*. *Animal Behaviour*, **55**, 1165–1179.

Kasperk, C., Helmboldt, A., Borcsok, I., Heuthe, S., Cloos, O., Niethard, F. & Ziegler, R. 1997. Skeletal site-dependent expression of the androgen receptor in human osteoblastic cell populations. *Calcified Tissue International*, **61**, 464–473.

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.

Lass, N. J. & Davies, M. 1976. An investigation of speaker height and weight identification. *Journal of the Acoustical Society of America*, **59**, 700–703.

Lass, N. J., Tecca, J. F., Mancuso, R. A. & Black, W. I. 1979. The effect of phonetic complexity on speaker race and sex identification. *Journal of Phonetics*, **7**, 105–118.

Lass, N. J., Hendricks, C. & Iturriaga, M. A. 1980. The consistency of listener judgements in speaker height and weight identification. *Journal of Phonetics*, **8**, 439–448.

Lieberman, P. 1984. *The Biology and Evolution of Language*. Cambridge, Massachusetts: Harvard University Press.

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

Linville, S. E. 1996. The sound of senescence. *Journal of Voice*, **10**, 190–200.

Locke, J. L. & Snow, C. 1997. Social influences on vocal learning in human and nonhuman primates. In: *Social Influences on Vocal Development* (Ed. by C. T. Snowdon & M. Hausberger), pp. 274–292. Cambridge: Cambridge University Press.

Maurer, D., Hess, M. & Gross, M. 1996. High-speed imaging of vocal fold vibrations and larynx movements within vocalizations of different vowels. *Annals of Otology, Rhinology and Laryngology*, **105**, 975–981.

Mitani, J. C. & Stuht, J. 1998. The evolution of nonhuman primate loud calls: acoustic adaptation for long-distance transmission. *Primates*, **39**, 171–182.

Moore, C. A. 1992. The correspondence of vocal tract resonance with volumes obtained from magnetic resonance images. *Journal of Speech and Hearing Research*, **35**, 1009–1023.

Morton, E. S. 1977. On the occurrence and significance of motivational structural rules in some bird and mammal sounds. *American Naturalist*, **111**, 855–869.

Mulac, A. & Giles, H. 1996. ‘You're only as old as you sound’: perceived vocal age and social meanings. *Health Communication*, **8**, 199–215.

Owren, M. J., Seyfarth, R. M. & Cheney, D. L. 1997. The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cynocephalus ursinus*): implications for production processes and functions. *Journal of the Acoustical Society of America*, **101,** 2951–2963.

Park, R. 1999. Star Wars. *Empire,* August, 90. London: EMAP.

Qvarnstrom, A. & Forsgren, E. 1998. Should females prefer dominant males? *Trends in Ecology and Evolution*, **13,** 498–501.

Schön Ybarra, M. 1995. A comparative approach to the nonhuman primate vocal tract: implications for sound production. In: *Current Topics in Primate Vocal Communications* (Ed. E. Zimmerman, J. D. Newman & U. Jürgens), pp. 185–198. New York: Plenum.

Walton, J. H. & Orlikoff, R. F. 1994. Speaker race identification from acoustic cues in the vocal signal. *Journal of Speech and Hearing Research*, **37,** 738–745.

Weary, D. M., Guilford, T. C. & Weisman, R. G. 1993. Peak shifts produced by correlated response to selection. *Evolution*, **47,** 280–290.

Wedekind, C. & Furi, S. 1997. Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity? *Proceedings of the Royal Society of London, Series B*, **264**, 1471–1479.

Wittenbrink, B., Hilton, J. L. & Gist, P. L. 1998. In search of similarity: Stereotypes as naı̈ve theories in social categorization. *Social Cognition*, **16,** 31–55.

Whiteside, S. P. 1998a. Identification of a speaker's sex: a study of vowels. *Perceptual and Motor Skills*, **86,** 579–584.

Whiteside, S. P. 1998b. Identification of a speaker's sex from synthesized vowels. *Perceptual and Motor Skills*, **87,** 595–600.