Human listeners attend to size information in domestic dog growls

Anna M. Taylor,[a] David Reby, and Karen McComb
Department of Psychology, University of Sussex, Falmer, BN1 9QH United Kingdom

The acoustic features of vocalizations have the potential to transmit information about the size of callers. Most acoustic studies have focused on intraspecific perceptual abilities, but here, the ability of humans to use growls to assess the size of adult domestic dogs was tested. In a first experiment, the formants of growls were shifted to create playback stimuli with different formant dispersions (Δf), simulating different vocal tract lengths within the natural range of variation. Mean fundamental frequency ($F0$) was left unchanged and treated as a covariate. In a second experiment, $F0$ was resynthesized and Δf was left unchanged. In both experiments Δf and $F0$ influenced how participants rated the size of stimuli. Lower formant and fundamental frequencies were rated as belonging to larger dogs. Crucially, when $F0$ was manipulated and Δf was natural, ratings were strongly correlated with the actual weight of the dogs, while when Δf was varied and $F0$ was natural, ratings were not related to the actual weight. Taken together, this suggests that participants relied more heavily on Δf, in accordance with the fact that formants are better predictors of body size than $F0$. © *2008 Acoustical Society of America.* [DOI: 10.1121/1.2896962]

I. INTRODUCTION

Although originally developed for human speech, the source-filter model of sound production has now been applied to the study of vocal communication in many animal species (see Fitch, 2000 and Munhall and Byrne, 2007, for overviews). The source-filter model presumes that there are two independent mechanisms involved in vocal production. The source signal is first produced when air expelled from the lungs passes through the vocal folds in the larynx and sends them into periodic vibration at a rate known as the fundamental frequency ($F0$). The signal then passes through the vocal tract, which, acting as a narrowband filter, enhances those frequencies associated with the resonances of the vocal tract while attenuating all others (Titze, 1994; also see Fitch, 2000). The amplification of the resonant frequencies causes broad energy bands in the spectrum, which are known as formants (Fant, 1960). $F0$ is thus determined by the length, mass, and tension of the vocal folds, while the values and spacing of formants are determined by the length of the vocal tract, leading to a negative correlation between formant dispersion (hereafter referred to as Δf) and vocal tract length in many mammals (see Fitch, 1997).

Formants and fundamental frequency are thus independent components in the vocal spectrum. This is important for human speech, which as a multidimensional form of communication can convey meanings as well as indexical information such as age, sex, individual identity, or geographical origin (see Rendall *et al.*, 2005). It is also significant for other mammal vocalizations that may need to communicate identity, affective state, or physical characteristics (e.g., Fitch and Hauser, 2002; Seyfarth and Cheney, 2003a; 2003b; Van-

noni and McElligot, 2006; Munhall and Byrne, 2007). Fundamental frequency, for example, has previously been linked to identity and emotional content in domestic dog barks (Yin, 2002; Yin and McCowan, 2004), while Δf has been linked to caller size because of its dependence on the vocal tract (Riede and Fitch, 1999). Indeed, anatomical constraints on the length of the vocal tract impose acoustic limits on formant dispersion, making formants an honest cue to size in many mammals (Fitch and Hauser, 2002; Reby and Mc-Comb, 2003a). In other words, larger animals produce lower and more closely spaced formants, whereas smaller animals produce higher, more widely spaced formants (macaques: Fitch, 1997; dogs: Riede and Fitch, 1999; red deer: Reby and McComb, 2003a; elephant seals: Sanvito *et al.*, 2007). Although there is some evidence for a correlation between vocal tract length and body size in humans (see Fitch and Giedd, 1999), the great flexibility of the human vocal tract generally means that Δf is a poor predictor of size in adult humans (e.g., González, 2004; Feinberg *et al.*, 2005; Rendall *et al.*, 2005).

Unlike the vocal tract, the vocal folds can develop independently of skeletal influences. As a result, correlations between $F0$ (which is a direct product of vocal fold length) and body size are generally weak or nonexistent in both humans and nonhuman mammals (Kunzel, 1989; McComb, 1991; Reby and McComb, 2003b; Rendall *et al.*, 2005). It should be noted, however, that when comparing individuals across age and sex, the vocal folds become a more reliable cue to body size. In those circumstances, associations between $F0$ and body size may be observed (Fischer *et al.*, 2002; Pfefferle and Fischer, 2006).

Many studies have now shown that animals naturally attend to formants in conspecific vocalizations (Sommers *et al.*, 1992; Fitch and Fritz, 2006; Charlton *et al.*, 2006) and that they are able to respond to size-related acoustic cues in a

---

[a]Author to whom correspondence should be addressed. Electronic mail: a.m.taylor@sussex.ac.uk

functionally significant way (Reby and McComb, 2003a; 2003b; Reby et al., 2005; Charlton et al., 2007). Furthermore, Ghazanfar et al. (2007) showed that rhesus monkeys were able to use formant cues to make accurate assessments of caller body size and age in conspecific vocalizations. Humans also attend to formants and formant-related parameters (e.g., Smith et al., 2005); however, due to the greater flexibility of the human vocal tract, it is often difficult to assess the body size of adult humans using only formants (e.g., Collins, 2000; González, 2004; Feinberg et al., 2005). It is also unclear whether listeners can assess size differences from Δf in the vocalizations of species other than their own. The domestic dog is an interesting candidate to investigate interspecific vocal communication, as its primary social companion is human rather than conspecific. Because of this distinct evolutionary background, most humans are likely to be exposed to dog vocalizations in their everyday life, and to thus have some degree of familiarity with these vocalizations.

To date there has been no attempt to investigate the intra- or interspecific perception of size information in dog vocalizations, despite it being a unique species in terms of the morphological variability found between breeds. No other has such a huge range of shapes and sizes (from under 2 kg to over 100 kg; see Sutter et al., 2007). Although in a previous study we found little evidence of context information encoded in the acoustic features of growls (growls recorded in a playful and in an aggressive context were compared; unpublished data), size information is likely to be present (see Riede and Fitch, 1999). The harsh broadband nature of growls makes them an ideal vocalization to investigate the role of formants in the transmission of size-related information by domestic dogs.

For an interspecific approach it is important to understand how humans normally perceive formants and how we interpret size information in speech. Several studies have demonstrated that humans are actually surprisingly bad at judging speaker size, mainly because we appear to base our assessment on acoustic features that are only weakly correlated with size in adult humans (Collins, 2000; González, 2004; Feinberg et al., 2005; Bruckert et al., 2006). Collins (2000) showed that women incorrectly associated lower frequency parameters with masculine traits such as larger size and body hair. Similarly, Feinberg et al. (2005) found that manipulations of Δf and $F0$ influenced the perceived attractiveness of male voices, with lower frequencies being rated as more attractive and belonging to larger men. Formant dispersion is thus perceptually salient to humans and ratings of size appear to be relatively consistent across different trials and individuals (Feinberg et al., 2005; Bruckert et al., 2006); however, they do not correspond to reality due to the poor relationship between formant and fundamental frequencies and body side in adult humans.

In a study of interspecific size perception solely based on a vocalization, we first investigate the relationship between formant dispersion and fundamental frequency in dog growls and caller body size. The predicted relationship between Δf and body size has previously been confirmed by Riede and Fitch (1999); however, the relationship between fundamental frequency and body size has not yet been examined in domestic dogs. Previous research on other species (Fitch, 1997; Fischer et al., 2002; Pfefferle and Fischer, 2006) suggests that this relationship will be very weak or even nonexistent. We then test the ability of humans to assess size information in dog growls with two psychoacoustic experiments that investigate on which acoustic parameters human listeners rely to attribute size to adult domestic dogs. In the first playback experiment, the formant dispersion of growls is modified but the fundamental frequency is left unchanged, whereas in the second playback experiment, fundamental frequency is modified and formant dispersion is left unchanged—this enables us to examine the independent effects of each of these acoustic features on size assessment. For the reasons previously discussed, it is expected that participants will assess the size of dogs using the same strategy as they use to assess human speaker size, relying on both formant dispersion and fundamental frequency.

II. METHODS

A. Dogs and recordings

The growls used in this study were recorded from 30 privately owned domestic dogs of 22 different breeds. The ages of the dogs ranged from 1 to 13 years; there were 13 females and 17 males, and all dogs except two (both Dobermanns) were neutered or sterilized. The dogs were recruited when their owners responded to a local advertisement and were recorded on their home territory in the presence of their owners between October 2005 and August 2006. The recordings were all made in the same context: Upon arrival at the house, the experimenter (A.M.T.) approached the dog while staring it in the eyes to elicit a defensive growling response (similar to the method used by Riede and Fitch, 1999). A Sennheiser MKH 416 directional microphone connected to a Marantz PMD670 digital audio recorder was used, with the microphone typically held at around one meter from the dog's head (although this varied a little as the dogs sometimes moved around). Recording levels were manually adjusted to avoid audio clipping. Finally, the dogs were weighed using PS250 veterinary floor scales.

B. Acoustic analyses

All audio manipulations and analyses were done with Praat acoustical analysis software, version 4.4.32 [by Boersma and Weenink (1996)]. The first high quality growl (defined as clearly audible with a well-defined beginning and ending and no background noise) from each dog was extracted in mono AIFF format with a 16 bit resolution and 44.1 kHz sampling rate. A Praat macro script, using the robust autocorrelation method described by Boersma (1993), was used to extract mean $F0$ from the whole duration of the signal, using a time step of 100 ms. The frequency range for the analysis was selected as 50–180 Hz. Two dogs in the present study were found to have a mean $F0$ outside the predicted range—both were Jack Russel terriers. Their mean $F0$ was thus separately extracted. Following Owren and Bernacki (1998), a second script relying on Linear Predictive Coding was used to measure formant frequencies in each

TABLE I. Biological and acoustic information collected for each dog.

Dog	Sex	Age (years)	Weight (kg)	Mean fundamental frequency (Hz)	Formant dispersion (Hz)
German shepherd	F	6	35.4	95	927
American cocker	M	3	12	148	1290
Basset hound	F	5	34.6	98	861
Cairn terrier	F	3.5	6.8	120	1285
Collie 1	M	3	18	117	944
Collie 2	F	2.5	12.5	130	1274
Dobermann 1	F	8	33	172	844
Dobermann 2	M	5	32	130	875
English cocker	F	2	5.8	127	1316
Flat-coated retriever	F	1	35	111	859
Jack Russell 1	F	6	4.2	204	1393
Jack Russell 2	M	7	6.8	193	1443
Jack Russell 3	M	2	8.5	114	1225
King Charles spaniel	M	4	8.3	189	1293
Labrador	M	10.5	26.5	124	1108
Lurcher 1	M	13	13.9	111	1143
Lurcher 2	M	2	17.6	134	993
Miniature dachshund	M	1	5	98	1433
Miniature schnauzer	F	2	6.6	185	1205
Old English sheepdog	M	12	35	110	948
Patterdale terrier	M	13	16	107	1186
Rhodesian ridgeback	M	6.5	25.1	178	871
Rottweiler	M	4	35.2	88	1143
Shi Tzu	M	8	7	119	1237
Springer spaniel 1	M	8	26.1	115	1153
Springer spaniel 2	F	1.5	20	149	1318
Staffordshire bullterrier 1	F	2	16.9	110	1435
Staffordshire bullterrier 2	F	6	17.7	138	1045
Staffordshire bullterrier 3	F	5	14	108	1291
Standard schnauzer	M	11	22.6	154	1049

growl. Where possible, five formants were measured; however, the number of formants varied from three to six. This is in line with previous observations in formant variations between different breeds of dogs (Riede and Fitch, 1999). Formant locations were also visually verified using a broadband spectrogram. Formant dispersion (Δf) was then computed using the linear regression method described in Reby and McComb (2003b). For each dog, there was thus a record of weight, mean $F0$, and Δf (see Table I for qualitative data).

C. Resynthesis

1. First playback experiment

Acoustic resyntheses were performed using pitch-synchronous overlap and add (PSOLA) based algorithms as coded into the "change gender" function in Praat. PSOLA algorithms enable signals to preserve their overall shape as target features are allometrically manipulated without affecting other acoustic parameters (see Goncharoff and Gries, 1998).

The durations of the stimuli were standardized to 2.8 s. Stimuli were never lengthened or shortened by more than 10% of their original duration to preserve the quality of the call. In addition, 1 s of silence was inserted before and after each stimulus. The use of PSOLA enables the manipulation of formants without changing $F0$, pitch envelope, amplitude

contour, or temporal structure; so only the formant frequencies were resynthesized in the first experiment while the other parameters were left natural. Resynthesized growls with formant dispersions of 844 (simulating a vocal tract length of 20.7 cm), 995 (17.6 cm), 1146 (15.3 cm), 1294 (13.5 cm), and 1443 Hz (12.1 cm) were created; 844 and 1443 Hz were the smallest and largest Δf observed in our natural population—the resynthesized stimuli thus represented a wide range within the natural variation of dog growls. Figure 1 shows spectrograms of one of the resynthesized stimuli, in which the dark energy bands of the formants are clearly visible. It is notable that formants in dog growls sometimes show dynamic modulation as seen in Fig. 1, and the occurrence and function of these still require further investigation. The resynthesis failed for ten growls, which were removed from the study. In total there were thus 100 resynthesized stimuli, originating from 20 individuals. Finally, all stimuli were double checked to make sure that no other acoustic parameters had been altered and their amplitudes were normalized to 75 dB (± 5 dB).

2. Second playback experiment

PSOLA was also applied to change the mean $F0$ of the growls without altering the formants or any other acoustic parameters. The way PSOLA affects $F0$ is by applying a
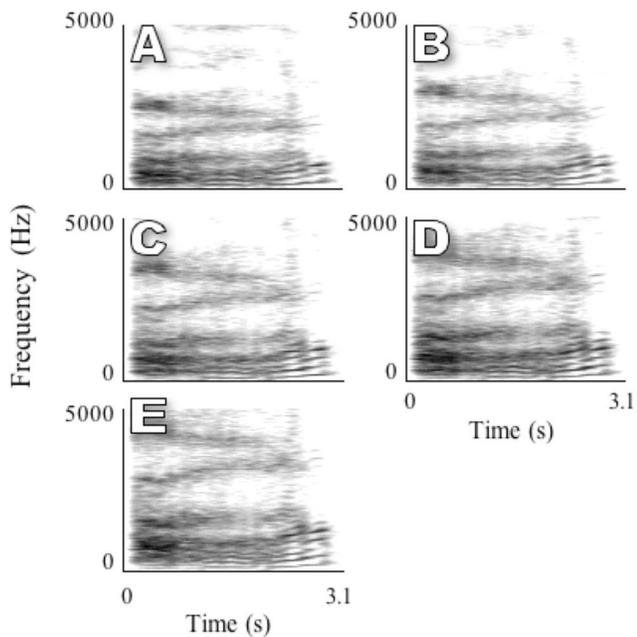
FIG. 1. Spectrograms (FFT window size=512 pts, window length=0.15 s) of the resynthesized stimuli created from the natural growl of a male Jack Russell terrier. The formant dispersions of the resynthesized stimuli are (A) 844, (B) 995, (C) 1146, (D) 1296, and (E) 1443 Hz. Note how the formants are shifted in comparison to the original signal, while all other acoustic parameters, including fundamental frequency, remain unchanged.

multiplicative transformation and thus controlling for relative frequency modulation in the signal (see Goncharoff and Gries, 1998). This preserves the natural shape of the signal. Each growl was resynthesized five times to give mean fundamental frequencies for our playback stimuli of 90, 119, 148, 177, and 200 Hz. The lowest and highest observed $F0$ in our sample were 90 and 200 Hz, respectively, so all the resynthesized stimuli were within the natural variation of dog growls. The resynthesis failed for 13 growls, which were removed from the study. In total there were thus 80 resynthesized stimuli originating from 17 individuals.

D. Playback protocol

1. First playback experiment

Thirty undergraduate students of the University of Sussex participated in the first playback experiment (25 females and five males; mean age=23 years). An MFC Praat experiment script running on an eMac (Apple Mac OS X, 1.4 GHz PowerPC G4) was used to run the playback. Participants were individually tested in the same lab. Testing was standardized by the use of the same location and equipment for all participants. They listened to each of the 100 stimuli in a randomized order through Koss Porta-pro headphones (frequency response: 15–25 000 Hz), and were then asked to rate each stimulus after the question "What is the size of this dog?" by clicking on a Likert scale from 1 (very small) to 7 (very big). The script automatically recorded the participant number, stimulus number, order of presentation, and size rating in a table file. Participants were then asked to state their previous dog experience from 1 (no experience) to 5 (regular contact with dogs). These data were transferred into an Excel
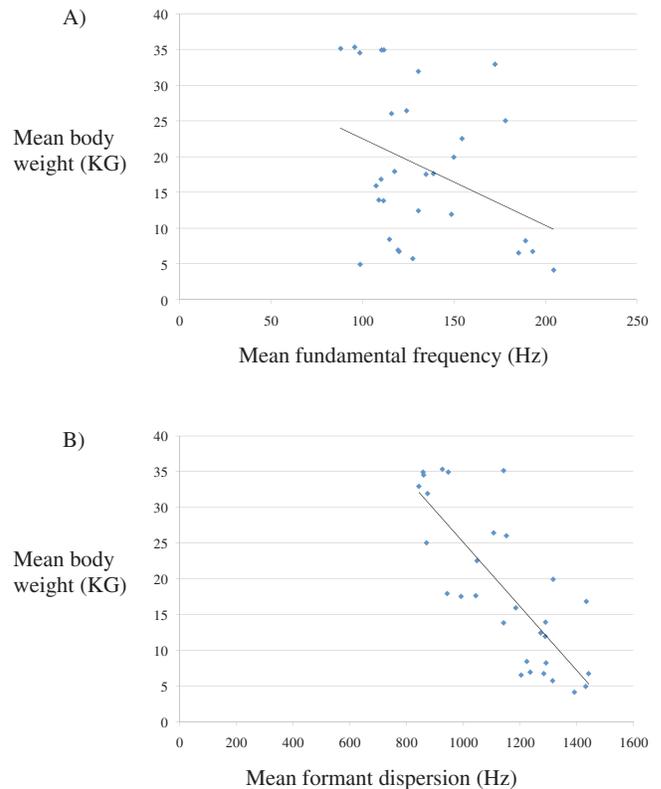

FIG. 2. (Color online) Single regression plots showing the relationship between the body weight of the dogs as a function of the two acoustic variables under study: mean (A) fundamental frequency and (B) formant dispersion.

file including the fundamental frequency of each stimulus (which varied between, but not within, dogs).

2. Second playback experiment

An identical protocol was used for the second playback experiment using a different group of 22 participants (19 females and three males; mean age=21). In this second experiment, participants listened to the playback stimuli within which only the fundamental frequency was resynthesized.

III. RESULTS

A. Acoustic analyses

Table I shows qualitative information about the 30 dogs that were recorded for this study, as well as the mean $F0$ and Δf measured in their growls. As we were interested in finding out to what extent these acoustic measures were predictive of caller weight, a multiple regression model was built entering Δf and mean $F0$ as predictors. The model found that, together, Δf and $F0$ accounted for $R^2_{(adj)}=66.3\%$ of the variation in weight ($F_{(2,27)}=29.48$, $p<0.001$). This was a stronger fit than a single regression model based solely on formant dispersion ($R^2_{(adj)}=62.3\%$, $F_{(1,28)}=48.84$, $p<0.001$) or on $F0$ ($R^2_{(adj)}=9\%$, $F_{(1,28)}=32.24$, $p=0.05$), demonstrating that when studying a species with a large morphological range, fundamental frequency can be weakly and negatively associated with body size. Plots of the single regressions are shown in Fig. 2 with body weight shown as a function of

acoustic parameters.

B. First playback experiment: Effect of formant dispersion on size ratings

A linear mixed model using a scaled identity covariance structure was used to analyze the data. The order of stimulus presentation and subject identity were entered as random factors in the model, Δf as a fixed factor and $F0$ and previous dog experience as covariates. The model was run with a maximum likelihood estimation and identified a significant main effect of Δf ($F_{(4,2300)}=56.28$, $p<0.001$) on size ratings. Ratings were also influenced by mean $F0$ ($F_{(1,2300)}=1027.07$, $p<0.001$) and previous dog experience ($F_{(4,2300)}=32.24$, $p<0.001$). These results indicate that recordings resynthesized to simulate a longer vocal tract (closely spaced formants) were rated as belonging to larger dogs and recordings resynthesized to simulate a shorter vocal tract (more widely spaced formants) were rated as belonging to smaller dogs. Examination of the residuals showed a homogeneous distribution, confirming a good fit of the linear mixed model (see Venables and Ripley, 2002). There was no significant interaction effect between manipulated Δf and natural $F0$ on size ratings ($p=0.107$).

C. Second playback experiment: Effect of fundamental frequency resynthesis on size ratings

Using the same statistical protocol as above, we found significant effects of resynthesized mean $F0$ ($F_{(4,3740)}=73.25$, $p<0.001$) on size ratings. Ratings were also influenced by natural Δf ($F_{(1,3740)}=49.11$, $p<0.001$) and previous experience ($F_{(1,3740)}=4.04$, $p=0.044$). Again, there was no significant interaction effect between natural Δf and manipulated $F0$ on size ratings ($p=0.120$). These results are in accord with the findings of the first playback experiment and confirm that humans rely on both formant dispersion and $F0$ to judge the size of dogs. Reliance on these cues appears to be influenced by previous dog experience, and indeed participants with a higher level of experience appear to show more reliance on Δf than participants with a lower levels of experience (see Sec. D).

D. Accuracy of the ratings

An investigation of the association between the mean size rating given to each dog (across all stimuli) and the actual weight of subjects revealed no correlation between these factors in the first playback, where Δf was resynthesized but $F0$ was natural ($r=0.35$, $p=0.13$). In the second playback, however, where $F0$ was resynthesized but formants were natural, there was a significant correlation between ratings and actual weight ($r=0.72$, $p=0.001$). This result held true within every level of experience, as shown in Table II. Interestingly, the strongest correlation between ratings and weight was found within the highest level of experience ($r=0.73$, $p=0.001$), suggesting that greater previous experience of dogs increases the ability to correctly assess their size. This result indicates that although both Δf and $F0$

TABLE II. Correlation coefficients between *participant size ratings* and *actual dog weight* within each level of previous dog experience.

Level of experience	First experiment		Second experiment	
	r	p	r	p
1[a]	…	…	…	…
2	0.29	0.22	0.67**	0.003
3	0.44*	0.05	0.64**	0.005
4	0.43	0.06	0.65**	0.004
5	0.38	0.10	0.73**	0.001

[a]No statistics were computed for this level of experience, as only one participant chose this option.

were used to assess the size of callers, participants relied more strongly on Δf and were thus not misled by manipulation of $F0$ in the second playback experiment.

IV. DISCUSSION

The acoustic analyses confirmed that formant dispersion in domestic dog growls is strongly predictive of caller weight, and thus growls have the potential to provide an honest cue to body size. This is in line with previous analyses in domestic dogs (Riede and Fitch, 1999) and other mammals (e.g., Fitch, 1997, Reby and McComb, 2003a). In addition, there was a weak, but significant, association between fundamental frequency and caller weight. Although correlations between $F0$ and body size have typically been found to be unreliable in adult mammals within a given species (McComb, 1991; Fitch, 1997; Fitch and Hauser 2002; Reby and McComb, 2003a), there is evidence that $F0$ can be indicative of body size in certain circumstances, for example when considered across age and/or gender (Pfefferle and Fischer, 2006). The great morphological variability between different dog breeds is likely to have led to broad variation in the length and mass of the vocal folds, and as such it is not surprising that there should be a link between $F0$ and caller weight in our mixed-breed sample.

Both Δf and $F0$ thus have the potential to convey body size in dogs, and in fact the results of our playback experiments showed that both these parameters had a significant effect on size ratings by participants. Stimuli resynthesized to simulate dogs with a longer vocal tract (closely spaced formants) were rated as larger than stimuli resynthesized to simulate dogs with a shorter vocal tract (widely spaced formants). Likewise, stimuli resynthesized to have a lower $F0$ were rated as larger than stimuli resynthesized to have a higher $F0$. The use of both Δf and $F0$ to make size assessments is not surprising, for it is already known that humans rely on both to judge speaker size within their own species. Because neither Δf nor $F0$ are reliably indicative of body size in adult humans (Kunzel, 1989; van Dommelen, 1993; van Dommelen and Moxness, 1995), this means that speaker size assessments are generally inaccurate (Collins, 2000; González, 2004; Feinberg et al., 2005).

Two plausible theories have been proposed for why human size assessments are so imprecise. First, the $F0$ of human speech is determined by the amount of testosterone present during puberty, at which point the larynx descends

J. Acoust. Soc. Am., Vol. 123, No. 5, May 2008

Taylor *et al.*: Size information in dog growls   2907

into its mature position and vocal fold length is established (Titze, 1994); this means that $F0$ is to some extent linked to age (mature versus immature). In addition, because of differences in male and female puberty, $F0$ is also a sexually dimorphic trait in humans. The $F0$ of adult males is thus approximately half that of adult females (Huber *et al.*, 1999). Developmental changes in $F0$ (within call types) have also been observed in some nonhuman animals such as red deer, where subadult stags have a higher $F0$ than adult stags (Fitch and Reby, 2001; Reby and McComb, 2003b). The perceived association between $F0$ and body size may thus stem from an overgeneralization of the existing association across age and sex (see Collins, 2000, and Feinberg *et al.*, 2005 for comprehensive discussions). Another possible reason for the inaccuracy of human size assessment could stem from our evolutionary history. $F0$ may have originally been a salient cue to maturity with formant cues providing additional information about the size of the mature male. Attendance to both $F0$ and Δf would thus have provided an evolutionary advantage before the development of the flexible vocal tract seen in modern humans (Collins, 2000). The overgeneralization and evolutionary advantage hypotheses are not mutually exclusive and may both have contributed to how humans assess speaker size, although further research into size perception mechanisms in humans will be necessary to determine which theory carries more weight.

We found no interaction effect of Δf and $F0$ on the ratings (in our playback protocol each set of participants was exposed to either resynthesized formants with natural $F0$ or resynthesized $F0$ with natural formants; a future study could investigate how formant dispersion and $F0$ affect size perception when both variables are resynthesized). Similarly, when assessing female mate preference based on red deer stag roars, researchers found no interaction effect when Δf and $F0$ were simultaneously varied (Charlton *et al.*, 2008). Red deer hinds, however, unlike humans, do not appear to use $F0$ for size estimation (Charlton *et al.*, 2006; 2007, McComb, 1991).

From our results, it is clear that participants used the same strategy to rate dog size as they use to rate the size of human speakers. Since both Δf and $F0$ are to some degree related to size across different breeds of dogs, it follows that reliance on both cues is justified when assessing the size of dogs. Formant dispersion was found to provide the strongest cue to size, as supported by the finding that size ratings from the second playback experiment were correlated with the actual weight of subjects. In this experiment, $F0$ was resynthesized but Δf was left unchanged, so it seems that participants were able to extract accurate size cues from the natural Δf without being misled by the resynthesized $F0$. In contrast the ratings of the first playback experiment (resynthesized Δf but unchanged $F0$) were not correlated with actual weights.

Overall, this shows that human listeners are able to accurately assess the size of domestic dogs, based on vocal signals alone. It also raises the question as to what (if any) significance human listeners attribute to $F0$. While there do not appear to be any acoustic differences between play growls and aggressive growls, it is possible that humans might perceive vocalizations with a lower $F0$ as more aggressive, in the same way that they perceived human speech with a lower $F0$ as more masculine in Collins' (2000) experiments. In addition, an effect of previous dog experience was found in both playback experiments—from this we propose that dog owners become more accurate at judging size as exposure to dogs teaches them not to over-rely on $F0$. Indeed, in the second experiment, the size ratings of the participants with the most previous dog experience were also the closest to reality, indicating that those participants were most confident about basing their ratings on the natural Δf and were less confused by the resynthesized pitch (see Table II).

This is the first evidence of interspecific size assessment based on vocal signals alone and suggests that fundamental acoustic principles that are common to mammals are relied on to make size assessments. This could have important evolutionary implications in terms of antipredator and/or prey selection strategies, as it would mean that signaling relationships are not confined to one species, and thus, as in the case of dogs and humans, interspecific size assessments based on acoustic signals alone can provide important information for listeners. Our findings also suggest, however, that humans with little experience of dogs may to some extent be vulnerable to manipulation by dog growls. Within a certain anatomical range, smaller dogs may be able to partly exaggerate their perceived size (or aggressiveness) by means of a lower $F0$. Capitalization on a propensity to over-relate $F0$ to body size becomes possible if smaller dogs are able to produce vocalizations with a lower $F0$ than would normally be expected for their size. Furthermore, it is also worth considering that a certain amount of vocal plasticity is anatomically possible in domestic dogs, who may be able to lengthen their vocal tracts by lowering the larynx to a small degree (see Fitch, 2000 for cine-radiographic evidence of vocal plasticity in several mammal species including the domestic dog). It is, however, not currently known whether dogs are able to use this potential to manipulate the acoustic features of their vocalizations.

In summary, human listeners used both Δf and mean $F0$ to assess the size of domestic dogs. As both of these acoustic parameters are reliably linked to caller weight, it follows that humans are able to accurately assess size in dogs by listening to their growls. Δf provided the more salient cue for size assessment; this was reflected by the fact that when $F0$ but not Δf was resynthesized, participants were still able to use the natural Δf to rate size and were not fooled by the resynthesized $F0$. Furthermore, experience played a role in both playbacks, and we suggest that the most plausible explanation for this is that dog owners learn by exposure not to over-rely on $F0$. People with a high level of previous dog experience might thus learn to more strongly focus on Δf than people who are less familiar with dogs (who may thus be more likely to over-rely on $F0$). Although learning may to some extent influence size assessment, this study nevertheless showed that humans were able to accurately rate the size of another species based solely on a vocal signal. Future research will reveal whether this ability exists in other interspecific signaling relationships or whether it is linked to the close evolutionary history between humans and dogs.

Boersma, P. (**1993**). "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio in a sampled sound," Proc. Inst. Phonetic Sci. Univ. Amsterdam **17**, 97–110.

Boersma, P., and Weenink, D. (**1996**). "Praat: A system for doing phonetics by computer (version 4.4.32)," Computer software, downloaded from http://www.praat.org/.

Bruckert, L., Liénard, J.-S., Lacroix, A., Kreutzer, M., and Leboucher, G. (**2006**). "Women use voice parameters to assess men's characteristics," Proc. R. Soc. London, Ser. B **273**, 83–89.

Charlton, B., Reby, D., and McComb, K. (**2006**). "Female perception of size-related formant shifts in red deer," Anim. Behav. **74**, 707–714.

Charlton, B., Reby, D., and McComb, K. (**2007**). "Female red deer prefer the roars of larger males," Cell. Mol. Biol. Lett. **3**, 382–385.

Charlton, B., Reby, D., and McComb, K. (**2008**). "Effect of combined source (F0) and filter (formant) variation on red deer hind responses to male roars," J. Acoust. Soc. Am. **123**, XXX.

Collins, S. A. (**2000**). "Men's voices and woman's choices," Anim. Behav. **60**, 773–780.

Fant, G. (**1960**). *Acoustic Theory of Speech Production* (Mouton, The Hague, The Netherlands).

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., and Perrett, D. I. (**2005**). "Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices," Anim. Behav. **69**, 561–568.

Fischer, J., Hammerschmidt, K., Cheney, D. L., and Seyfarth, R. M. (**2002**). "Acoustic features of male baboon loud calls: Influences of context, age and individuality," J. Acoust. Soc. Am. **111**, 1465–1474.

Fitch, W. T. (**1997**). "Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques," J. Acoust. Soc. Am. **102**, 1213–1222.

Fitch, W. T. (**2000**). "The phonetic potential of nonhuman vocal tracts: Comparative cineradiographic observations of vocalizing animals," Phonetica **57**, 205–218.

Fitch, W. T., and Fritz, J. B. (**2006**). "Rhesus macaques spontaneously perceive formants in conspecific vocalizations," J. Acoust. Soc. Am. **120**, 2132–2141.

Fitch, W. T., and Giedd, J. (**1999**). "Morphology and development of the human vocal tract: A study using magnetic resonance imaging," J. Acoust. Soc. Am. **106**, 1511–1522.

Fitch, W. T., and Hauser, M. D. (**2002**). "Unpacking 'honesty:' vertebrate vocal production and the evolution of acoustic signals," in *Acoustic Communication*, edited by A. M. Simmons, R. R. Fay, and A. N. Popper (Springer-Verlag, New York), pp. 65–137.

Fitch, W. T., and Reby, D. (**2001**). "The descended larynx is not uniquely human," Proc. R. Soc. London, Ser. B **268**, 1669–1675.

Ghazanfar, A. A., Turesson, H. K., Maier, J. X., van Dinther, R., Patterson, R. D., and Logothetis, N. H. (**2007**). "Vocal tract resonances as indexical cues in the rhesus monkeys," Curr. Biol. **17**, 425–430.

Goncharoff, V., and Gries, P. (**1998**). "An algorithm for accurately marking pitch pulses in speech signals," *Proceedings of the International Conference on Signal and Image Processing 1998*. Retrieved from http://ieeexplore.ieee.org/iel5/9711/30650/01415093.pdf on 08/27/2007.

González, J. (**2004**). "Formant frequencies and body size of speaker: A weak relationship in adult humans," J. Phonetics **32**, 277–287.

Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A., and Johnson, K. (**1999**). "Formants of children, women and man: The effects of vocal intensity variation," J. Acoust. Soc. Am. **106**, 1532–1542.

Kunzel, H. (**1989**). "How well does average fundamental frequency correlate with speaker height and weight?," Phonetica **46**, 117–125.

McComb, K. E. (**1991**). "Female choice for high roaring rates in red deer, *Cervus elaphus*," Anim. Behav. **41**, 79–88.

Munhall, K. G., and Byrne, S. K. (**2007**). "Animal communication: Big talkers and small talk," Curr. Biol. **17**, 247–249.

Owren, M. J., and Bernacki, R. H. (**1998**) "Applying Linear Predictive Coding (LPC) to frequency-spectrum analysis of animal acoustic signals," in *Animal Acoustic Communication*, edited by S. L. Hopp, M. J. Owren, and C. S. Evans, pp. 299–346.

Pfefferle, D., and Fischer, J. (**2006**). "Sounds and size: Identification of acoustic variables that reflect body size in hamadryas baboons, *Papio hamadryas*," Anim. Behav. **72**, 43–51.

Reby, D., and McComb, K. (**2003a**). "Anatomical constraints generate honesty: Acoustic cues to age and weight in the roars of red deer stags," Anim. Behav. **65**, 519–530.

Reby, D., and McComb, K. (**2003b**). "Vocal communication and reproduction in deer," Adv. Stud. Behav. **33**, 231–264.

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., and Clutton-Brock, T. (**2005**). "Red deer stags use formants as assessment cues during intrasexual agonistic interactions," Proc. R. Soc. London, Ser. B **272**, 941–947.

Rendall, D., Kollias, S., and Ney, C., and Lloyd, P. (**2005**). "Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry," J. Acoust. Soc. Am. **117**, 944–955.

Riede, T., and Fitch, T. (**1999**). "Vocal tract length and acoustics of vocalization in the domestic dog (Canis familiaris)," J. Exp. Biol. **202**, 2859–2867.

Sanvito, S., Galimberti, F., and Miller, E. H. (**2007**). "Vocal signalling of male southern elephant seals is honest but imprecise," Anim. Behav. **73**, 287–299.

Seyfarth, R. M., and Cheney, D. L. (**2003a**). "Meaning and emotion in animal vocalizations," Ann. N.Y. Acad. Sci. **1000**, 32–55.

Seyfarth, R. M., and Cheney, D. L. (**2003b**). "Signalers and receivers in animal communication," Annu. Rev. Psychol. **54**, 145–173.

Smith, D. R. R., Patterson, R. D., Turner, R., Kawahara, H., and Irino, T. (**2005**). "The processing and perception of size information in speech sounds," J. Acoust. Soc. Am. **117**, 305–318.

Sommers, M. S., Moody, D. B., Prosen, C. A., and Stebbins, W. C. (**1992**). "Formant frequency discrimination by Japanese macaques (*Macaca fuscata*)," J. Acoust. Soc. Am. **91**, 3499–3510.

Sutter, N. B., Bustamante, C. D., Chase, K., Gray, M. M., Zhao, K., Zhu, L., Padhukasahasram, B., Karlins, E., Davis, S., Jones, P. G., Quignon, P., Johnson, G. S., Parker, H. G., Fretwell, N., Mosher, D. S., Lawler, D. F., Satyaraj, E., Nordborg, M., Lark, K. G., Wayne, R. K., and Ostrander, E. A. (**2007**). "A single IGF1 allele is a major determinant of small size in dogs," Science **316**, 112–115.

Titze, I. R. (**1994**). *Principles of Vocal Production* (Prentice–Hall, Englewood Cliffs, N.J.).

van Dommelen, W. A. (**1993**). "Speaker height and weight identification: Reevaluation of some old data," J. Phonetics **21**, 337–341.

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

Vannoni, E., and McElligot, A. G. (**2006**). "Individual acoustic variation in fallow deer (*Dama dama*), Common and harsh groans: A source-filter theory perspective," Ethology **113**, 223–234.

Venables, W. N., and Ripley, B. D. (**2002**). *Modern Applied Statistics with S*, 2nd ed. (Springer-Verlag, New York).

Yin, S. (**2002**). "A new perspective on barking in dogs (*Canis familiaris*)," J. Comp. Psych. **116**, 189–193.

Yin, S., and McCowan, B. (**2004**). "Barking in domestic dogs: Context specificity and individual identification," Anim. Behav. **68**, 343–355.