

REVIEW

The contribution of source–filter theory to mammal vocal communication research

A. M. Taylor & D. Reby

Mammal Vocal Communication and Cognition Research, School of Psychology, University of Sussex, Brighton, UK

Keywords

vocal communication; acoustic signals; source–filter theory; vocal production.

CorrespondenceAnna M. Taylor, Mammal Vocal Communication and Cognition Research, School of Psychology, University of Sussex, Brighton BN1 9QH, UK.
Email: a.m.taylor@sussex.ac.uk

Editor: Steven Le Comber

Received 13 July 2009; revised 25 September 2009; accepted 28 September 2009

doi:10.1111/j.1469-7998.2009.00661.x

Abstract

The field of animal vocal communication has benefited greatly from improved understanding of vocal production mechanisms and specifically from the generalization of the source–filter theory of speech production to non-human mammals. The application of the source–filter theory has enabled researchers to decompose the acoustic structure of vocal signals according to their mode of production and thereby to predict the acoustic variation that is caused by anatomical or physiological attributes of the caller. The source–filter theory states that vocal signals result from a two-stage production, with the glottal wave generated in the larynx (the source), being subsequently filtered in the supralaryngeal vocal tract (the filter). This theory predicts that independent indexical information such as body size, weight, age and sex can be contained in both the glottal wave (mostly characterized by its fundamental frequency), and the spectral envelope of the radiated vocalization (mostly characterized by the vocal tract resonances or formant frequencies). Additionally, physiological fluctuations in emotional or motivational state have been found to influence the acoustic characteristics of signals in a reliable and predictable manner that is perceptually available to receivers. While animal vocalizations contain some dynamic attributes, their static attributes are sufficient to provide an effective means of acoustic individual discrimination both within and across call types. In this paper, we draw together a wealth of experimental work conducted within the source–filter framework over the last decade and we review how such experiments have elucidated the communicative value of animal vocalizations.

Introduction

Understanding communication systems is essential to the study of animal behaviour and ecology, as the progression of interactions between individuals is mediated by visual, olfactory and vocal signals (Bradbury & Vehrencamp, 1998). In particular vocal signals have been found to play a crucial role in determining the outcome of intra- and intersexual competition and to mediate agonistic or affiliative interactions between individuals (Owings & Morton, 1998; Fitch & Hauser, 2002). In mammals, early research on communication focused primarily on the more conspicuous features of acoustic signalling such as call occurrence, calling rate and loudness and signaller/receiver interactions (Clutton-Brock & Albon, 1979; McComb, 1991; Owings & Morton, 1998; McElligott & Hayden, 1999), providing valuable insights into our understanding of the function and evolution of sound signals. In recent years, vocal communication research has benefited from the application of the ‘source–filter theory’ (Fant, 1960; Titze, 1994), a framework initially developed for the study of human

speech, which fits the requirements of a model linking vocal production, acoustic structure and functional discrimination/perception.

The aim of the present paper is thus to highlight how the source–filter theory has contributed to the current state of knowledge on vocal production mechanisms and its impact on animal vocal communication. Coupled with the development of modern digital techniques of signal analysis, the source–filter theory has enabled researchers to develop specific hypotheses within a testable framework. It is possible to investigate the detailed structure and variation of the acoustic parameters that compose vocalizations in relation to their origin, to describe and quantify their variation using clearly defined terminology, and finally to test their function using playback experiments (Fitch, 1997; Reby & McComb, 2003a,b; Charlton, Reby & McComb, 2007b; Charlton, McComb & Reby, 2008a; Koren & Geffen, 2009). We here present the key advances in the application of this framework. In the first section, we introduce the source–filter framework itself, define the acoustic parameters according

to their mode of production and make broad predictions about the information they are likely to encode in mammal signals. In subsequent sections, we review the impact of this approach on different aspects of mammal vocal signals categorized according to the nature and function of the information encoded in signals: static cues to fitness (second section), motivational and referential cues (third section) and cues to individual identity (fourth section). In each section, we present two types of studies: correlational approaches that examine the covariation of acoustic parameters with traits or events, and experimental approaches, where playbacks of acoustic stimuli are used to examine the perceptual and/or functional relevance of these parameters.

The source–filter framework

Speech scientists have determined that the production of the voiced signals that form human speech follows a two-stage process known as the ‘source–filter theory of voice production’ (Fant, 1960; Singh & Singh, 1976; Titze, 1994). According to this theory, the production of voiced signals involves independent contributions from different parts of the vocal apparatus, specifically the ‘source’, which includes the larynx and all sub-laryngeal and laryngeal structures, and the ‘filter’ or ‘vocal tract’, which is defined as the tube linking the larynx to the openings (mouth and nose) from which sound radiates into the environment (Titze, 1994). It should be noted that several studies preceding the explicit application of source–filter theory to non-human mammals nevertheless fall conceptually into the source–filter framework (e.g. Masataka, 1994). The explicit conceptualization and generalization of the source–filter theory to vertebrates originated in bioacoustics research in the 1990s (Hauser, 1993; Newton-Fischer *et al.*, 1993; Fitch, 1994, 1997; Solomon, Luschei & Liu, 1995; Owren, Seyfarth & Cheney, 1997; Rendall, Owren & Rodman, 1998; Rendall *et al.*, 1999; Riede & Fitch, 1999; also see earlier discussions by Lieberman, 1975, 1984) and is based on the principle that the vocal production apparatus is fundamentally similar across mammalian species, including humans (Titze, 1994; Fitch & Giedd, 1999). The source–filter model has also been generalized to avian species (ring doves: Beckers, Suthers & ten Cate, 2003; Fletcher *et al.*, 2004; Elemans, Zaccarelli & Herzl, 2008; parrots: Beckers, Nelson & Suthers, 2004). Indeed, the avian syrinx performs a ‘source’ role similar to the larynx, and the avian trachea provides a ‘filter’ akin to the mammalian vocal tract (Fitch, 1999). While in this review we focus on mammals, we make occasional references to the avian literature for comparative purposes.

The ‘source’

In humans, it has long been established that production of voiced sounds commences in the larynx, a cartilaginous structure that is situated low in the throat where the tract of the pharynx splits into the trachea and the oesophagus.

The largest structures of the larynx are the thyroid cartilage (which is attached to the hyoid bone by the thyrohyoid membrane) and the cricoid cartilage (which forms the inferior wall of the larynx and attaches to the top of the trachea). The vocal folds are located at the superior border of this cricoid cartilage. They are attached at the back to the arytenoid cartilages and at the front to the thyroid cartilage. The vocal folds themselves consist of three layers: muscle, vocal ligament and the epithelium. They are sometimes referred to as ‘vocal cords’, however, the term ‘vocal folds’ is preferred when discussing mammals as is it more anatomically correct (Titze, 1994; Fitch, 2006). Together with the spacing between them, the vocal folds form the glottis, where voiced sounds are generated. As air from the lungs forces its way through the closed glottis, the vocal folds are pushed apart. Biomechanical forces cause the vocal folds to snap shut again, and this sequence of opening and closing of the glottis causes a cyclic variation in air pressure across the larynx. Earlier accounts of vocal production stated that vocal fold vibration was predominantly driven by Bernoulli forces building up from sub-glottal pressure (van den Berg, Zantema & Doornenbal, 1957; Fant, 1960; Lieberman, 1977); however, systems of mechanical vibration invoked by Bernoulli forces are subject to dampening out, resulting in a gradual decrease in mechanical activity (Fung, 1981; Chan & Titze, 2006). A better understanding of tissue biomechanics has enabled researchers to determine that the continuous energy provided by the airflow from the lungs as it passes through the vocal folds creates a self-sustaining system of ‘flow-induced oscillation’. In such a system no additional mechanical forces are necessary to maintain a continuous rate of vibration (see Chan & Titze, 2006 for a detailed account of flow induced oscillation). The resulting waveform constitutes the source signal or glottal wave. While the vocal anatomy of all non-human mammals is fundamentally the same, most non-human mammals have a more elevated laryngeal position than humans with the larynx attached to the skull in a static position at the back of the oral cavity (Fig. 1).

The rate of opening and closing of the glottis determines the fundamental frequency (henceforth ‘F0’) of the glottal wave, also sometimes referred to as the glottal pulse rate. In human speech, F0 is the main factor determining the perceived pitch of a voice (however, it should be noted that the term ‘pitch’ is essentially perceptual and is better avoided when describing acoustic variation in vocal signals). F0 is determined primarily by the length and mass of the vocal folds: longer and heavier vocal folds vibrate at a slower rate than smaller vocal folds (Titze, 1994; Fitch, 1997). In humans, these properties can to a certain extent be manipulated by the muscles that form the vocal folds: flexion or relaxation of the cricothyroid or thyroarytenoid muscles control the lengthening and shortening of the vocal folds, and flexion or relaxation of the cricoarytenoid muscles control adduction and abduction of the vocal folds (Fink, 1975; Hardcastle, 1976). Other characteristics of the source signal include tempo, duration and amplitude contour, all of which are controlled by sophisticated muscular

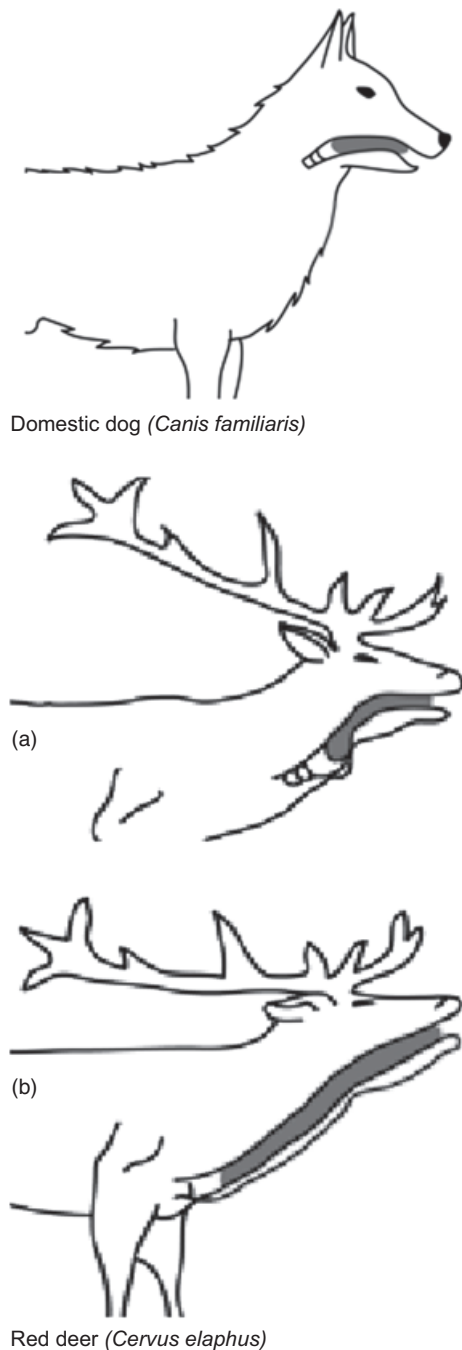


Figure 1 Position of the larynx in domestic dogs *Canis familiaris* and red deer *Cervus elaphus* stags (domestic dog adapted from Fitch, 2000c; red deer stags reproduced with permission from Fitch & Reby, 2001). Domestic dog default 'high' laryngeal resting position (most common in non-human mammals). Red deer (a) 'low' laryngeal resting position (uncommon in non-human mammals); (b) fully descended larynx during male roars.

interactions and changes in airflow or sub-glottal pressure (Titze, 1994). Generally speaking, in both humans and non-human animals, the acoustic characteristics of the glottal wave are not reliably related to body size, because the organs that produce them are soft and unconstrained by skeletal structures (Fitch, 1997, 2000b). Source characteristics can thus vary between and within vocalizations from the same caller either on a volitional (intonation in human speech: Ohala, 1984; Banse & Scherer, 1996; frequency modulation in bats: Bastian & Schmidt, 2008) or on an involuntarily basis (emotional expression in humans: Ohala, 1996; Aubergé & Cathiard, 2003; affective state in baboons: Rendall, 2003b; stress in pigs: Döpjan *et al.*, 2008).

While the source signal is generally periodic, many recent studies report the presence of non-periodic elements (or 'non-linear phenomena') in the source component of mammalian vocal signals. Although in humans, non-linear phenomena can be related to speech pathologies (e.g. Hirano, 1981), in many non-human animals they form part of the normal vocal communication system. Examples of non-linear phenomena include subharmonics (additional harmonics visible in the spectrum beneath F₀; African wild dogs: Wilden *et al.*, 1998; chimpanzees: Riede, Owren & Arcadi, 2004), biphonation (two independent F₀; African wild dogs: Wilden *et al.*, 1998; H. S. Webster *et al.*, unpubl. data; chimpanzees: Riede *et al.*, 2004; dholes: Volodina *et al.*, 2006) and deterministic chaos (broadband signals with no particular harmonics; African wild dogs: Wilden *et al.*, 1998; chimpanzees: Riede *et al.*, 2004, red deer: Reby & McComb, 2003a,b). Bifurcations between linear and non-linear events are also often observed in species presenting non-linear phenomena (Wilden *et al.*, 1998; Fitch, Neubauer & Herzog, 2002; Tokuda *et al.*, 2002; Riede *et al.*, 2004). Despite improvements in our understanding of the production process and role of non-linear phenomena in human speech (Titze, 2008), their place in animal communication systems is not yet well defined, although several hypotheses are discussed in the literature (Wilden *et al.*, 1998; Fitch *et al.*, 2002; Tokuda *et al.*, 2002; Riede *et al.*, 2004, 2005; Riede, Arcadi & Owren, 2007).

The 'filter'

The second stage of the source-filter theory is the filtering process that takes place in the vocal tract between the production of the signal at the source and its external radiation. The vocal tract consists of all the air cavities between the larynx and the opening of the mouth and/or nostrils through which the source signal travels before it radiates into the environment (Fig. 1). The vocal tract acts as a bank of bandpass filters, selectively dampening and/or enhancing specific ranges of frequencies from the source signal, corresponding to the resonant properties of its physical structures. The resonant frequencies form spectral peaks called formants (from the Latin *formare*, to shape; Fant, 1960; Titze, 1994).

In humans, the two largest cavities of the vocal tract are the pharynx and the mouth (Titze, 1994). Sophisticated vertical and horizontal movements of the tongue and lower jaw in the pharynx and the mouth influence the resonant properties of the vocal tract, thereby affecting the relative frequency position of formants, and particularly that of the lower formants (Fant, 1960; Lieberman, Klatt & Wilson, 1969; Hauser, Evans & Marler, 1993; Titze, 1994). Modulation of the lower formants of the voice spectrum results in the production of the different phonemes we perceive as vowels (Fant, 1960; Titze, 1994). In non-human animals, the vocal tract is usually not as flexible and thus its resonant properties are often static and more predictable (Fitch, 1994; 2000a,b, 2002). In particular the length of the vocal tract is directly reflected in the formants of many animal vocal signals (Fitch, 1997).

We have so far stressed that an important assumption of source-filter theory lies in the independence of source and filter, enabling researchers to relate specific acoustic parameters to their mechanism of production. However, it should be noted that in some circumstances interactions between source and filter components have been observed when the source or the filter influences or interferes with the output of the other (Titze, 2008). The contribution of source-filter interactions to the diversity of mammal vocal signals remains to be fully investigated.

Production constraints and cues to physical attributes

Animals use vocal communication to mediate crucial interactions such as sexual competition, territorial maintenance, partner or parent/young recognition and coordination of defence against predators (Owings & Morton, 1998). The outcome of many of these interactions depends on the physical attributes of individuals, such as their body size, physical condition, age or sex (Schmidt-Nielsen, 1975; Peters, 1986; Andersson, 1994). A comprehensive discussion of how acoustic signals may have the potential to provide accurate and reliable information about the physical attributes of individuals is given in a seminal paper by Fitch & Hauser (2002). Here we update the notion of 'honest signalling' (Fitch & Hauser, 2002) with a range of empirical tests conducted within the source-filter framework. Acoustic cues to physical attributes are often referred to as 'indexical' (Ghazanfar *et al.*, 2007), for they provide receivers with reliable information on specific attributes (e.g. size) of callers (Fitch, 1997; Fitch & Reby, 2001; Maynard Smith & Harper, 2003). While the role of such cues often originates in the physical relationship that ties them to the dimension they express, selection pressures have often led to senders being able to alter their values to some extent (Ohala, 1984; Maynard Smith & Harper, 2003), an aspect that is discussed in more detail below. Understanding the origin, nature and function of such acoustic indexical cues is therefore one of the most active areas of current vocal communication research.

Indices generated at the source

Initially, studies of animal vocal signals tended to focus on understanding the control and variability of F0 (Cohen & Fox, 1976; Tembrock, 1976; Morton, 1977; August & Anderson, 1987; Masataka, 1994). This focus is likely to have been a reflection of the salience of 'pitch' to human speakers and listeners (Ohala, 1984). F0 is perceptually identifiable by non-specialists and it is easy to measure on spectrograms or oscillograms (see Boersma, 1993). Moreover, F0 is highly variable within and between calls both across individuals and across species (Morton, 1977; August & Anderson, 1987; Hauser, 1993; Yin, 2002; Reby & McComb, 2003a,b; Rendall *et al.*, 2005). In an influential paper based on a comparative study of vocalizations used in agonistic displays in a range of mammalian and avian species, Morton (1977) suggested that audible frequency differences in vocalizations reflect ritualized signalling: animals with aggressive motivation produce low-pitched, broadband vocalizations (such as growls and hisses), while animals with a friendly or submissive motivation produce high-pitched vocalizations (such as whimpers and whines). This theory, known as Morton's motivation-structural code, is based on the observation across several species that aggressive and dominant animals seek to project (both visually and acoustically) a larger impression of body size whereas friendly or submissive animals seek to project a smaller impression of body size (Morton, 1977; Ohala, 1984; Owings & Morton, 1998). It is well documented that larger sized individuals are at an advantage over smaller individuals during agonistic encounters, and individuals benefit from avoiding escalation of unmatched encounters, due to the great risk of injury or even death caused by fighting (Schmidt-Nielsen, 1975; Clutton-Brock & Albon, 1979; Peters, 1986).

We have seen that the F0 of vocal signals is determined by the physical properties and adjustments of the vocal folds. However, due to their soft tissue anatomy, the growth of the vocal folds is not stringently constrained by the body size of an individual (Fitch, 1997, 2000c). A good illustration of this can be seen in the comparison of subspecies of *cervidae*. While adult Scottish red deer stags (weighing 160–250 kg) produce calls with a mean F0 of 112 Hz (Reby & McComb, 2003a), the smallest red deer subspecies (the Corsican deer, weighing *c.* 80 kg) produces calls with a mean F0 of 34 Hz and the largest red deer subspecies (the wapiti, weighing *c.* 400 kg) produces calls with a F0 up to 2080 Hz (Feighny, Williamson & Clarke, 2006; Riede & Titze, 2008), despite having a vocal fold length of about 3 cm, which would normally be expected to produce a F0 of *c.* 50 Hz (Riede & Titze, 2008). While to date it is unclear how the wapiti is able to produce such a high F0 (vocal fold elasticity alone cannot explain this extreme divergence from biomechanical predictions: Riede & Titze, 2008), this example provides a clear illustration of the independence of F0 from body size and even in this case from vocal fold length.

Across age and sex categories, possibly due to age-related vocal fold growth and sexual dimorphism, F0 can be

correlated with caller body size (e.g. in both baboons and humans, males are larger than females and also have a lower F₀; Rendall *et al.*, 2005; Pfefferle & Fischer, 2006; Puts, Gaulin & Verdolini, 2006). The same is true of some species in which unusually large morphological variations exist across individuals that in all other ways have identical developmental and reproductive patterns (e.g. different breeds of domestic dogs; Taylor, Reby & McComb, 2008). However, within most species and between members of same age or sex categories, there is ample evidence for a high level of independence between F₀ and body size (baboons: Rendall *et al.*, 2005; Japanese macaques: Masataka, 1994; red deer: Reby & McComb, 2003a; rhesus macaques: Fitch, 1997; but see Pfefferle & Fischer, 2006). In general, muscular control of the vocal folds means that F₀ has the potential to be modulated as the tension, length and mass of the vibrating segment is manipulated. Indeed, the range of variation of F₀ within individuals is often comparable to the variation between individuals (red deer: Reby & McComb, 2003a, dogs: Yin, 2002). This dynamicity means that F₀ may serve as a reliable indicator of other characteristics that are relevant to resource holding potential and mate selection, such as age, sex and dominance rank (humans: Fitch & Giedd, 1999; Rendall *et al.*, 2005; baboons: Rendall *et al.*, 2005; Pfefferle & Fischer, 2006; fallow deer: Vannoni & McElligott, 2008; red deer: Reby & McComb, 2003a,b). The type of information encoded in F₀ varies between species; thus in fallow deer males a lower F₀ is linked to high dominance status and higher reproductive success (Vannoni & McElligott, 2008), whereas conversely in red deer stags, F₀ is positively correlated with reproductive success (Reby & McComb, 2003a) and recent playbacks have shown that hinds prefer roars with a high F₀ (D. Reby *et al.*, unpubl. data).

In humans, one of the main drivers of vocal fold development is testosterone (Titze, 1994; Fitch & Giedd, 1999; Evans *et al.*, 2008): the testosterone increase during male puberty causes thickening and lengthening of the vocal folds, resulting in a decrease in F₀ of about 50% in comparison to same-aged women (in contrast, the body size variation between adult men and women is *c.* 20%; Fitch & Giedd, 1999). The F₀ of human male voices presents diurnal variations in line with the circulating level of sex hormones (Evans *et al.*, 2008) and similarly, cyclical changes in the F₀ of female voices have been found to be linked to the hormonal variations controlling the menstrual cycle (Abitbol, Abitbol & Abitbol, 1999; Caruso *et al.*, 2000; Pipitone & Gallup, 2008). Similar hormonally induced physiological changes could be at the basis of F₀ changes observed when non-human mammals reach sexual maturity, with sub-adults generally producing a higher F₀ than mature males (baboons: Fischer *et al.*, 2002; red deer: Reby & McComb, 2003a). In red deer, the vocal folds continue to grow in length after the animal itself has stopped growing, resulting in a strong correlation between vocal fold length and age throughout the lifetime of individuals (Reby & McComb, 2003b). When considering individuals across the whole developmental spectrum, F₀ thus appears to co-vary with

age (specifically with sexual maturity; baboons: Fischer *et al.*, 2002; red deer: Reby & McComb, 2003a) and sex (baboons: Rendall *et al.*, 2005; Pfefferle & Fischer, 2006; fallow deer: Vannoni & McElligott, 2008; red deer: Reby & McComb, 2003a).

Indices generated by the filter

Realizing the importance of filter-induced variation in animal vocalizations has been one of the most exciting recent developments in bioacoustics. Unlike the vocal folds, the vocal tract cannot grow independently of the rest of the body for its development is anatomically constrained by skeletal structures (Fitch, 2000b,c). The vocal tract length is thus directly dependent on body size. Investigations have confirmed a strong negative correlation between vocal tract length and body size (domestic dogs X-rays: Riede & Fitch, 1999; red deer dissections: Fitch & Reby, 2001; rhesus macaque radiographs: Fitch, 1997). This means that, unlike F₀, formant frequencies have the potential to provide accurate or 'honest' information about the caller (Fitch, 1997, 2000c; Fitch & Reby, 2001; Fitch & Hauser, 2002; Reby & McComb, 2003b).

The overall spacing between formants appears to play the greatest role in providing an acoustic correlate of caller size. This relationship is quantified under the term 'formant dispersion' (Titze, 1994; Fitch, 1997; Reby & McComb, 2003a), literally referring to the pattern of dispersion of formants in the spectrum of the call. A direct negative correlation between formant dispersion and body size (Japanese macaques: Fitch, 1997; red deer: Reby & McComb, 2003a; domestic dogs: Riede & Fitch, 1999; Taylor *et al.*, 2008; pandas: Charlton, Zhang & Snyder, 2009) has been confirmed in many species. Figure 2 illustrates the relationship between the formant dispersion calculated from growl vocalizations in 30 domestic dogs of different breeds and their respective body weight.

When the importance of formant dispersion as a size code was first identified, it was calculated as the 'average distance

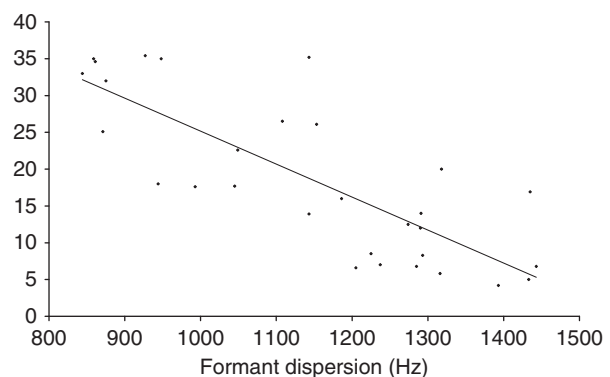


Figure 2 Body weight (kg) of 30 domestic dogs *Canis familiaris* plotted as a function of their formant dispersion (Hz) calculated from growl vocalizations (adapted from Taylor *et al.*, 2008). $R^2=62.3\%$, $P<0.001$, $n=30$.

between each adjacent pair of formants' (Fitch, 1997, p. 1216). An alternative method of calculation is to run a linear regression plotting observed formant locations against predicted formant spacing. When the vocal tract is modelled as a straight uniform tube that is closed at one end and open at the other, the spacing between any two successive formants (Δf) can be approximated as a constant, and formant frequencies can be plotted as $F_i = \Delta f \frac{(2i-1)}{2}$, as illustrated in Figure 3 (Reby & McComb, 2003a). Regardless of which method of calculation is used, formant dispersion can be used to estimate vocal tract length by the equation $VTL = \frac{c}{2(\Delta f)}$, where c is the speed of sound in air approximated as 350 m s^{-1} and Δf is the formant dispersion (Titze, 1994; Fitch, 1997).

The observation that formant dispersion has the potential to provide an accurate acoustic representation of caller body size (Fitch, 1997; Reby & McComb, 2003a; Taylor *et al.*, 2008) has led to a series of studies investigating whether receivers use size-related acoustic variation to assess callers. Spontaneous discrimination of size-related formant variation has been demonstrated in several species using habituation-discrimination paradigms (rhesus macaque: Fitch & Fritz, 2006; whooping crane: Fitch & Kelley, 2000) and the behavioural consequences of formant discrimination have been investigated (red deer: Reby *et al.*, 2005; Charlton, Reby & McComb, 2007a,b; Charlton *et al.*, 2008a,b; dogs: A. M. Taylor, D. Reby & K. McComb, unpubl. data). Moreover, rhesus monkeys are able to associate smaller formant dispersions with pictures of larger (mature) conspecifics and wider formant dispersions with pictures of smaller (immature) individuals (Ghazanfar *et al.*, 2007), demonstrating an intermodal (auditory to visual) understanding of size. In humans, formant shifts as small as 7% are picked up by listeners (Smith & Patterson, 2005; Rendall, Vokey & Nemeth, 2007), and can influence how a

speaker is perceived by other men and women in terms of weight, height, masculinity and dominance (Collins, 2000; Bruckert *et al.*, 2006; Puts *et al.*, 2007; Rendall *et al.*, 2007).

In some species, callers have evolved anatomical adaptations that enable them to alter the relationship between body size and formant frequency dispersion in their vocal signals. Both red and fallow deer show an anatomical peculiarity that was previously believed to be unique to humans: instead of the larynx resting in an elevated position at the back of the oral cavity as seen in most non-human mammals, the larynges of male red and fallow deer rest in an unusually low position in the neck (Fig. 1; red deer: Fitch & Reby, 2001; fallow deer: McElligott, Birrer & Vannoni, 2006). This causes the vocal tracts of these animals to be longer than would normally be expected for their size. Consequently, their vocalizations contain lower formant dispersions relatively to other species lacking this anatomical innovation, in effect resulting in the projection of a relatively exaggerated impression of their body size. As illustrated in Fig. 4, the larynges of male red and fallow deer can be retracted even further into the throat during the production of mating calls, allowing further vocal tract elongation (Fitch & Reby, 2001; McElligott *et al.*, 2006). Laryngeal retraction is made possible by the evolution of a highly elastic thyrohyoid membrane linking the larynx to the hyoid apparatus and strong sternothyroid and sternohyoid muscles that pull the larynx down the throat toward the sternum (Fitch & Reby, 2001). As the sternothyroid and sternohyoid muscles are attached to the sternum, the larynx cannot be pulled lower than the sternum, putting an anatomical limitation on laryngeal retraction and thereby maintaining the proximate honesty of this signal (Fitch & Reby, 2001; Fitch & Hauser, 2002). A similar anatomical adaptation enabling laryngeal retraction during mating calls has also been observed in Mongolian gazelles (Frey *et al.*, 2008). Moreover, as already

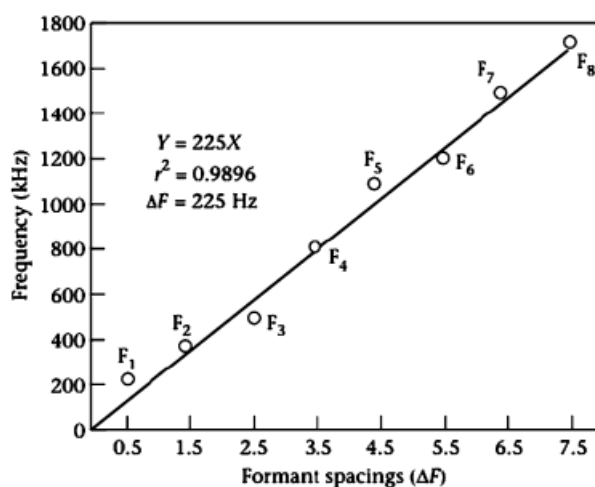


Figure 3 Using the linear regression method to estimate formant dispersion, referred to here as 'formant spacing'. In this example, the formant dispersion is 225 Hz, corresponding to a vocal tract length of 0.78 m (reproduced with permission from Reby & McComb, 2003a).

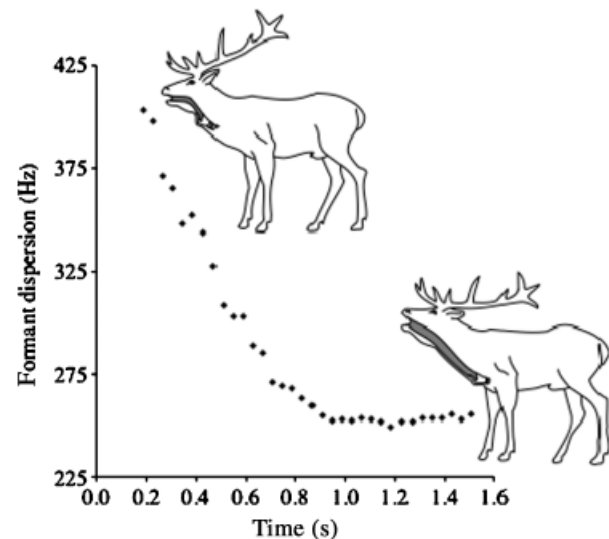


Figure 4 Relationship between laryngeal retraction and change in formant dispersion during roaring in red deer *Cervus elaphus* stags (reproduced with permission from Fitch & Reby, 2001).

noted by Fitch (2000b, 2006), several other behavioural and anatomical adaptations may be involved in acoustic size exaggeration. For example, male saiga antelopes are able to increase the length of their vocal tract while producing mating calls by means of a specific vocal posture involving a strongly tensed and extended trunk (Volodin, Volodina & Efremova, 2009). Furthermore some species possess a pronounced proboscis, elongating the nasal region of the vocal tract and potentially influencing the spacing of formant frequencies (elephant seals: Sanvito, Galimberti & Miller, 2007). Similarly, black and white colobus monkeys have evolved a subhyoid airsac that is inflated to act as an additional resonator during roars, thereby lowering their formants in comparison to what would normally be observed for animals of the same body size (Harris *et al.*, 2006; also see Riede *et al.*, 2008 for an experimental test of the effect of laryngeal airsacs on formant frequencies). On a comparative note, at least 60 species of birds possess elongated tracheas, and the evolution of this has been discussed in the context of the size exaggeration hypothesis (see detailed review by Fitch, 1999).

Vocalizations are an integral part of male competitive signalling (Bradbury & Vehrencamp, 1998; Owings & Morton, 1998). The size-related variation in formants can thus provide receivers with valuable information about potential competitors, and enable functional decisions about whether or not to escalate an agonistic interaction with another individual, based on the assessment of the caller's body size relative to that of the receiver (red deer: Fitch & Reby, 2001; Reby *et al.*, 2005; fallow deer: McElligott *et al.*, 2006; domestic dogs: A. M. Taylor, D. Reby & K. McComb, 2009b). As well as affecting interactions linked to male–male competition, acoustic size exaggeration (or maximization) also appears to play an important role in mate choice (Charlton, 2008). A recent mate choice experiment demonstrated that red deer hinds preferentially approached a loudspeaker emitting stag roars resynthesized to have more closely spaced formants mimicking a larger male, over a loudspeaker emitting roars resynthesized with more widely spaced formants mimicking a smaller male (Charlton *et al.*, 2007b). This suggests that sexual selection for anatomical adaptations mediating acoustic size exaggeration may be a driving factor in the evolution of these production mechanisms (Fitch & Reby, 2001; Charlton, 2008). It has been hypothesized that the lowered resting position of the larynx in humans may have evolved through similar selection pressures, predating the development of speech (Ohala, 2000; Fitch & Reby, 2001; Fitch, 2002).

There is compelling evidence that formant information is also perceived across species, presumably because the fundamental similarities across mammal vocal production systems have led to comparable similarities in the perception of acoustic signals. Several animals have been trained to discriminate vowel-like sounds using operant conditioning techniques (Chacma baboons: Hienz & Brady, 1988; Chinchilla: Burdick & Miller 1975; domestic dogs: Baru, 1975; Japanese macaques: Sinnott, 1989; Sinnott & Kreiter, 1991; Sommers *et al.*, 1992). Using resynthesized formants,

researchers furthermore demonstrated that human listeners were able to reliably rate the size of domestic dogs based on an acoustic signal alone (Taylor *et al.*, 2008), providing direct evidence for interspecific perception and assessment of size-related variation in formant frequencies. The use of formants as indices of body size may be widespread in mammals with potential implications for interspecific interactions such as eavesdropping predator/prey contexts.

Finally, Fitch (1997) notes that reliability of size information in formants is dependent on the quality of the source signal. Formants are perceptually easier to discriminate in harsh, broadband calls (such as grunts, groans or growls) than in high F₀, tonal calls with wide inter-harmonic intervals and little inter-harmonic energy. The impact of some source characteristics on formant perceptibility is little investigated and remains an area of interest for future empirical work.

Affective and referential information

In the previous section, we have shown how acoustic signals are frequently dependent on static physical attributes, and also how anatomical or behavioural adaptations may effectively provide a means of vocal control. In the context of social interactions, the significance of vocal signals may go beyond the encoding of caller attributes and may provide a secondary level of information relating to the current motivational or emotional state of individuals (Ohala, 1984). Additionally, in a limited number of mammals, the acoustic structure of signals can be arranged in a predictable manner, so as to provide functionally referential information (Hauser, 1993, 1998), defined here as signals providing reliable stimulus-specific information about an external object or event (Macedonia 1993) that are produced only in an appropriate context (Macedonia 1993; Hauser, 1993, 1998; Evans, 1997). The investigation of affective basis and referential content in animal vocalizations is highly relevant in the light of understanding the evolution of human speech and how meaning has become encoded in phonetic variability, bringing the source–filter theory to the centre of this topic (Fitch, 2000a, 2002; Ohala, 2000; Slocombe & Zuberbühler, 2005).

Dynamic encoding at the source

In many species, there are significant differences between calls recorded in different social situations (baboons: Owren *et al.*, 1997; Rendall *et al.*, 1999; Seyfarth & Cheney, 2003a,b). This is true both between call types (i.e. specific types of vocalizations occur consistently in specific contexts; Morton, 1977) and within call types, where the acoustic structure of call varies according to context (domestic dogs barks: Yin, 2002; Yin & McCowan, 2004). Indeed, several characteristics of F₀ (such as mean F₀, peak F₀ and F₀ modulation) have been linked to the context in which calls are emitted (baboons: Fischer *et al.*, 2002; domestic dogs: Yin, 2002; Taylor *et al.*, 2009a; pandas: Charlton *et al.*, submitted; wapiti: Feighny *et al.*, 2006; also see Ohala,

1984). Classification methods such as discriminant function analysis are useful in confirming the acoustic categorization of vocalizations emitted in different contexts. For example, Yin (2002) found that domestic dogs barks occurred on a graded scale, showing a continuum of acoustic gradations on several frequency parameters depending on the situation in which they were emitted. It was confirmed that barks could be statistically divided into different context-specific subsets on the basis of the co-variation of their peak, mean fundamental frequency, duration and inter-bark interval (Yin & McCowan, 2004). These parameters furthermore enabled human listeners to reliably categorize barks in function of their recording context (Pongrácz *et al.*, 2005).

Dynamic changes in F0 providing cues to affective state are most likely mediated by changes in physiological arousal such as rate of respiration or muscular (cricothyroid) tension in the vocal folds (Scherer, 1986; Titze, 1994; Hauser, 2000; Bachorowski & Owren, 2008). Generally speaking, the motivational information provided by F0 fits the framework of the motivation-structural rules and frequency code theory: thus, the barks of domestic dogs recorded in an aggressive context have been found to have a significantly lower F0 than barks recorded in a playful setting (Yin, 2002; Yin & McCowan, 2004; Pongrácz *et al.*, 2005; Taylor *et al.*, 2009). Similarly, wapiti bugle calls emitted in aggressive contexts are lower in frequency (both F0 and formants) than bugle calls emitted during non-aggressive interactions (Feighny *et al.*, 2006). In fallow deer, males with a low F0 have a higher reproductive success, and likewise human men with a low F0 report more sexual partners, and are more attractive to women, than males with higher F0 (fallow deer: Vannoni & McElligott, 2008; humans: Puts 2005; Apicella, Feinberg & Marlowe, 2007; Apicella & Feinberg, 2009). However, the effect of motivational state on F0 does not always follow a predictable direction. For example, male baboons with a high dominance status produce calls with a higher F0 than lower ranked males (Fischer *et al.*, 2004), presumably because they have a high reproductive and territorial motivation and are in a higher state of physiological arousal. Similarly, red deer stags with a higher F0 are known to have a greater reproductive success than stags with lower F0 (Reby & McComb, 2003b). It thus seems prudent to propose that variations in F0 are species-specific and should be documented across several species before generalized assumptions across species can be made.

Another dimension of the source implicated in the communication of motivational state is calling rate. Calling rate can be linked to rate of respiration, and typically provides immediate information about the current condition or motivation of an individual (red deer: Clutton-Brock & Albon, 1979; McComb, 1991). During the rutting season, fallow deer bucks call at a rate of 3000 groans h^{-1} – groaning in this species appears to be aimed at other males by advertising a measure of fighting motivation rather than at attracting females (McElligott & Hayden, 1999, 2001). Interestingly, fallow deer bucks may perform less laryngeal retraction in favour of maintaining a high groaning rate, as the latter plays

a more important role in this species (Vannoni, Torriani & McElligott, 2005). In contrast, red deer stags, who retract the larynx to some degree for virtually all roars, are able to sustain a roaring rate of ‘only’ around 400–500 roars h^{-1} (Clutton-Brock & Albon, 1979; McComb, 1991). This trade-off between calling rate (indicating physical condition and fitness) and laryngeal retraction (indicating body size, as well as fitness) may be due to the dual role of roaring in intra-sexual competition and mate attraction which may differ between these two species.

Finally, calling rate and call duration may also be communicative of urgency (Blumstein & Armitage, 1997; Manser, 2001; Seyfarth & Cheney, 2003a,b; Furrer & Manser, 2009). In general, higher calling rates, combined with longer vocalizations are indicative of urgent contexts, whereas slower calling rates with shorter vocalizations are typical of more relaxed contexts (Rendall *et al.*, 1999; Seyfarth & Cheney, 2003a,b; Fischer *et al.*, 2004). In domestic dogs, higher barking rates are observed when barks are recorded in aggressive situations (Pongrácz *et al.*, 2005), and both barks and growls are significantly longer when produced in aggressive contexts (Yin, 2002; Taylor *et al.*, 2009). Similarly, baboons grunting rate increases with heightened arousal (Rendall *et al.*, 1999), and in both marmots and meerkats, chirping rate and call duration and intensity are increased when a predator threat is more imminent (Blumstein & Armitage, 1997; Manser, 2001; Manser, Seyfarth & Cheney, 2002). Rapidly pulsing sounds activate the sympathetic nervous system, increasing physiological arousal and creating an internal sensation of urgency (McConnell & Baylis, 1985; McConnell, 1990), and in fact longer signals may be perceived as louder than shorter calls (McConnell & Baylis, 1985; Le Roux, Jackson & Cherry, 2001). ‘Loudness’ and variations in intensity or amplitude contour are also dependent on sub-glottal pressure, which tends to increase with heightened arousal and/or motivation (bison: Wyman *et al.*, 2008; non-human primates: Seyfarth & Cheney, 2003a,b); this is most likely due to increases in respiration-related airflow.

Dynamic encoding in the filter

Because of the overall reliability of formant frequencies as an acoustic correlate of body size, small variations in formant dispersion may have a secondary function of signalling motivational state. Effectively, the signalling of body size can become a ritualized advertisement of emotional or motivational state (Ohala, 1984; also see Morton’s, 1977 motivation-structural rules). In several species, callers have been observed to retract the lips in positive situations or in encounters where it is beneficial for them to appease another individual (such as a smile or fear grin; canines: Fox, 1970; humans: Drahotá, Costall & Reddy, 2008) and to protrude the lips during socially stressful or agonistic encounters where it is beneficial to appear larger or more dominant (baboons: Harris *et al.*, 2006; wolves: Fox, 1970). Moreover, Ohala (1984) has proposed that the vocal gestures associated with different emotional/motivational

states may have driven the evolution of facial expressions, a theoretical hypothesis that has received support from both observational and empirical studies (Harris *et al.*, 2006; Chuenwattanapranithi *et al.*, 2008; Drahota *et al.*, 2008). The interaction between emotional/motivational state, acoustic output and facial expression is a largely unexplored branch of vocal communication (also see Ohala, 1984; Aubergé & Cathiard, 2003; Chuenwattanapranithi *et al.*, 2008; Drahota *et al.*, 2008) and further research in non-human mammals is required to determine the full extent and limitations of this phenomenon.

Complex control of filter components has also been observed for the encoding of context-specific information. Indeed, some non-human primate species produce acoustically distinct alarm calls for different classes of predators (Barbary macaques: Diana monkeys: Zuberbühler, Cheney & Seyfarth, 1999; Zuberbühler, 2002; vervet monkeys: Seyfarth, Cheney & Marler, 1980; Owren, 1990a; Owren & Bernacki, 1998; rhesus monkeys: Hauser, 1998; also see meerkats for a non-primate example: Manser, 2001; Manser *et al.*, 2002), and other group members are able to respond in functionally appropriate ways to these calls (vervet monkeys look up or hide under bushes in response to an 'eagle' alarm call, but scan their surroundings and head for the trees in response to a 'leopard' alarm call: Seyfarth *et al.*, 1980; Seyfarth & Cheney, 1990). Functionally referential calls, at least in some primate species, appear to evolve along a continuum whereby purely reflexive/affective calls come under more volitional control (Macedonia & Evans, 1993; Evans, 1997). Thus innate distress calls may have become more and more specific throughout evolution, driven by audience effects and the receiver comprehension, and culminating in voluntarily alarm calling (Sherman, 1977; Cheney & Seyfarth, 1985; Seyfarth & Cheney, 2003a,b). In acoustic terms, using both natural and resynthesized stimuli, it has been found that the discrimination between 'snake' and 'eagle' alarm calls by conspecifics in vervet monkeys is most reliable when made using spectral cues, even though temporal and fundamental frequency cues also vary between the two calls (Owren, 1990a,b; Owren & Bernacki, 1998; Seyfarth & Cheney, 2003a,b). In fact, the active modulation of the first two formants during vocalizations appears to play the greatest role in referential communication, with deviations from what would be expected of a uniform vocal tract ranging from 23% for F1 to 60% for F2 (Riede & Zuberbühler, 2003; Riede *et al.*, 2005, 2008). This is perhaps not surprising as F1 and F2 are dependent on those parts of the vocal tract that have the most potential for volitional manipulation. Rudimentary modulation of the first two formants is reminiscent of the process seen in the acoustic differentiation of vowel sounds in human speech, as the vocal tract is manipulated in order to filter the source signal specifically to encode external events (Fant, 1960; Lieberman & Blumstein, 1988). These results support the hypothesis that the shaping of spectral patterns in alarm calls is likely to have evolved specifically for communicative reasons, and may be paramount in the transition from purely affective calls (all mammals) to functionally referential calls

(some non-human primates), and ultimately to intentionally referential calls (humans) (see Evans, 1997).

Inter-play between source and filter

We have seen that source and filter components can provide varying levels of affective and functionally referential information in many mammalian species. In human speech, the combination of source and filter characteristics is vital for language as both intonation and semantic content are necessary for successful communication (Lieberman & Blumstein, 1988). In non-human mammals, the potential inter-play and communicative effects of interactions between source and filter is less well understood (but see Charlton *et al.*, 2008b), although recent research has shown that hyrax songs simultaneously encode body weight, size, current condition, hierarchical status and current hormonal state of the singer (Koren & Geffen, 2009). It is likely that several levels of information may be similarly present within the signals of other mammals, and this largely unexplored branch of animal vocal communication merits further investigation.

Cues to individual identity in acoustic signals

Acoustic distinctiveness both at an individual and at a group level (Rendall, Rodman & Emond, 1996) has been shown to be essential for many different species; for example to maintain group cohesiveness when foraging in dense undergrowth or in otherwise reduced visibility, for the recognition of familiar versus unfamiliar individuals in neighbouring or overlapping territories, or for the identification of kin in large groups or in species where mother and offspring are not constantly together (Trivers, 1971; Halliday, 1983). Moreover, acoustic familiarity may play a role in inter-sexual interactions where the familiarity of males may be an indication of their investment in sexual displays toward females (Zimmermann & Lerch, 1993; Reby *et al.*, 2001). Acoustic variation may originate in the source or in the filter, and understanding their relative contributions to individuality, and how selection pressures have differentially affected these contributions, has greatly assisted our understanding of the form and function of different types of vocal identity cues.

Source-induced individual variability

Several source characteristics have been implicated in individual distinctiveness, including amplitude contour, harmonic structure including harmonic-to-noise ratio and the presence of subharmonics and temporal features such as signal tempo and duration, (baboons: Rendall, 2003a; domestic dogs: Yin, 2002; chimpanzees: Riede *et al.*, 2004; coyotes, dogs: Riede *et al.*, 2005; fur seals: Charrier, Mathévon & Jouventin, 2003b; rhesus monkeys: Rendall *et al.*, 1998; roe deer: Reby *et al.*, 1999). These aspects of the glottal wave may also contribute to the voice distinctiveness

across call types, affecting the voice's timbre or 'harshness' independently of the filter (Riede *et al.*, 2004, 2005).

For many mammals however, the greatest contribution of the source to individual variation appears to be based on the dynamic modulation of F0 (baboons: Owren *et al.*, 1997; bottlenose dolphins: Janik 2000; fallow deer: Torriani, Vannoni & McElligott, 2006; wolves: Palacios, Font & Márquez, 2007), which is furthermore linked to the existence of 'vocal signatures' (fur seals: Charrier *et al.*, 2003*b*; wolves: Palacios *et al.*, 2007; bottlenose dolphins: Janik, Sayigh & Wells, 2006). Signature calls essentially appear to serve a similar identifying purpose to 'names' in human interactions, although it should be noted that in human speech, names are words composed of phonemes produced by the manipulation of formants (Lieberman & Blumstein, 1988). F0 mediated acoustic distinctiveness has been identified between group members but also across individuals within a same group, so that all members within the group produce a recognizable call that is distinct from the signature call of other groups (hyenas: Holekamp *et al.*, 1999; pigtail macaques: Gouzoules, Gouzoules & Marler, 1995). Moreover, F0 appears to be especially important for kin recognition in many species, specifically for reuniting mother and offspring. Different selection pressures have led to variations in acoustic kin recognition strategies. In species using a hider strategy where the young remain hidden in the undergrowth after birth (see Langbein & Putman, 1992), there is a selection pressure for young to be silent to avoid detection by predators. Because the offspring are mobile and may change hiding places, acoustic recognition of the dam is essential to maintain maternal care in these species (Fisher, Blomberg & Owens, 2002). Fallow deer fawns thus only leave their hiding place in response to the recognition of the distinctive fundamental frequency of their dam's call, while the dam does not recognize the call of the fawn (Vannoni *et al.*, 2005; Torriani *et al.*, 2006). On the other hand, in species using a follower strategy and in species where offspring are mixed with other same-aged offspring, mutual recognition between mother and young is vital for the survival of offspring (banded mongoose: Muller & Manser, 2008; northern fur seals: Insley, 2001; sheep: Searby & Jouventin, 2003). Recognition is likely to have evolved via different selection pressures on mother and young: for young animals, non-recognition of their mother may lead to death, whereas for the mother, non-recognition of their young may lead to the loss of one breeding season (Trivers, 1971). These differential pressures mean that acoustic recognition between mother and offspring may be asymmetrical (Insley, 2001). Thus in fur seals, pups attend to the harmonic structure and tempo of female calls to identify their mother (Charrier *et al.*, 2003*b*), while mothers appear to attend to the properties of the energy spectrum (frequency modulation and amplitude contour) to identify their pup (Charrier, Mathevon & Jouventin, 2002). Playback experiments in which the harmonic structure of maternal calls has been modified have unambiguously shown that this manipulation irrevocably impairs recognition in fur seal pups (Charrier *et al.*, 2003*b*).

Filter-induced individual variability

For some mammals and specifically for several primate species including humans, filter components play a substantial role in the acoustic distinctiveness of individuals (baboons: Owren *et al.*, 1997; red deer: Reby *et al.*, 1998; rhesus monkeys: Rendall *et al.*, 1998; also see Fischer *et al.*, 2001). This appears to be primarily an acoustic consequence of morphological individuality. For example, Rendall (2003*a,b*) used statistical algorithms to show that baboon grunts were individually distinct across several acoustic parameters (notably tempo, F0 and formant structure), but that formants provided the highest degree of differentiation between individuals due to the less reliable, dynamic nature of tempo and F0.

While most studies have focused on individual differences occurring within the same call types, there is some evidence that in some non-human mammals, individuals have idiosyncratic voices, like human speakers. Indeed, just as our voice is recognizable independently of the phoneme pronounced, in mammal calls the individuality of voice also has the potential to remain recognizable across call types, due to the physiological basis of the filtering process. This has been demonstrated in several species, giving credence to the concept of an 'individual voice' (baboons: Rendall *et al.*, 1998; red deer: Reby *et al.*, 2006). Red deer can be accurately individually identified across several call types (harsh roars, roars, chase barks and barks) due to inter-individual acoustic variation that most likely reflects individual differences in the morphology of the vocal tract (Reby *et al.*, 2006). Similarly, rhesus monkeys retain distinctiveness across coos, grunts and noisy screams (Rendall *et al.*, 1998), although the authors also note that individual distinctiveness across call types can sometimes be hampered by the broad structural differences between the calls.

Co-variation of source and filter

Several studies have now highlighted the importance of the inter-play of source and filter components for reliable identification of a caller (fallow deer: Reby *et al.*, 1998; Vannoni & McElligott, 2007; rhesus monkeys: Rendall *et al.*, 1998). In the case of mother–young recognition there is an interesting asymmetry: while adults do not typically vary in size, their offspring have growing bodies. Given the direct dependence of filter-related components on skeletal size, these acoustic parameters are expected to change allometrically in line with the physical development and growth of the offspring. Conversely, the relative independence of the source-related components from physical attributes means that they are potentially less subject to the developmental changes of the caller. In several pinniped species, it has been shown that mothers have long-term recognition of both the immature and adult vocalizations of their offspring from previous years (Insley, 2000; Charrier, Mathevon & Jouventin, 2003*a*). It would thus be of interest for future research to investigate the differential variation in source and filter characteristics throughout the

lifetime of individuals and how this co-variation might definitely affect individual distinctiveness in adults versus immature animals. A point of interest that emerges from the literature is the apparent evolutionary convergence of bleat vocalizations. Bleats are stereotypical plaintive vocalizations that occur across several unrelated species in the context of individual recognition (seal pups: Schustermann & van Parijs, 2003; sheep: Searby & Jouventin, 2003; Sèbe *et al.*, 2008). This highlights a potentially promising area for future research, as it seems likely that their acoustic characteristics are particularly favourable to individual and specifically mother–young recognition.

Conclusion

In this review, we have shown that the source–filter theory goes a long way in predicting, identifying and explaining the functional content of mammal acoustic signals and their evolution. We have presented many examples of how source- and filter-related components in vocalizations play independent roles in the transmission of different types of information, and specifically how they can function as indices of attributes such as body size, weight, age or sex. Attending to such cues, and using them to assess the physical attributes and condition both of potential competitors and mates, can have important implications for the reproductive opportunities and survival of receivers. In addition to static attributes, cues to transient qualities such as emotional or motivation state and dynamic qualities such as reproductive status or dominance rank can also be advertised in the source and filter components of vocal signals. Moreover, there is growing evidence that in some primate species callers are able to produce vocalizations containing information about events or objects in the external world encoded in their source and filter-related component characteristics. Finally, we have discussed how the inter-individual variation in anatomy/physiology reflected in the acoustic structure of vocal signals can lead to voice differences between individuals, and more specifically, how identity information can be given by frequency or amplitude contours, as is observed in the identify-

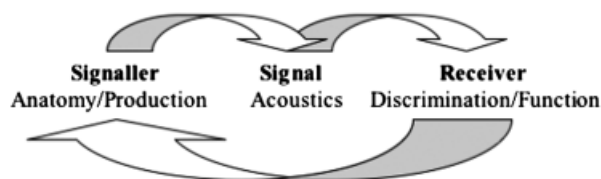


Figure 5 Evolutionary feedback loop linking voice production, acoustic output and function. The vocal apparatus and its operation determine the structure of the acoustic signal. The receiver has access to information encoded in the signal and makes functional decisions based on this information. These functional decisions drive selection of the signal at the level of its production, and therefore variations in the acoustic content of the signal is constrained by limits on variation of the vocal apparatus and its operation. Understanding the production mechanism of animal vocalizations is thereby crucial, for it is at this level that characteristics accessible to receivers, and therefore to selection, are determined.

ing whistles, and more generally in the ‘vocal signatures’, of several species.

In conclusion, this review has highlighted the important contributions of the source–filter paradigm to understanding mammal vocal communication. Understanding call production mechanisms has enabled the development of a testable framework for the investigation of the origin and function of signals. This conclusion is illustrated in Fig. 5, which provides an overview of the evolutionary feedback loop linking production mechanisms to the acoustic structure of signals and the ultimate effect this has on the perception by, and behaviour of receivers.

Acknowledgements

Many thanks to Karen McComb and Ben Charlton for their helpful comments on earlier versions of the paper, and to Alan McElligott for his support throughout the writing process. Thanks also to the contributions of Tim Halliday and one anonymous referee. Funded by a BBSRC studentship to the first author.

References

- Abitbol, J., Abitbol, P. & Abitbol, B. (1999). Sex hormones and the female voice. *J. Voice* **13**, 424–446.
- Andersson, M. (1994). *Sexual selection*. Princeton: Princeton University Press.
- Apicella, C.L. & Feinberg, D.R. (2009). Voice pitch alters mate-choice-relevant perception in hunter-gatherers. *Proc. Roy. Soc. Lond. Ser. B* **276**, 1077–1082.
- Apicella, C.L., Feinberg, D.R. & Marlowe, F.W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biol. Lett.* **3**, 682–684.
- Aubergé, V. & Cathiard, M. (2003). Can we hear the prosody of smile? *Speech Comm.* **40**, 87–97.
- August, P. & Anderson, J. (1987). Mammal sounds and motivation-structural rules: a test of the hypothesis. *Am. Soc. Mammal.* **68**, 1–9.
- Bachorowski, J. & Owren, M.J. (2008). Vocal expressions of emotion. In *Handbook of emotions*, 3rd edn: 196–210. Lewis, M., Haviland-Jones, J.M. & Feldman, L. (Eds). New York: Guilford Press.
- Banse, R. & Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* **70**, 614–636.
- Baru, A. (1975). Discrimination of synthesised vowels [a] and [i] with varying parameters (fundamental frequency, intensity, duration and number of formants) in dogs. In *Auditory analyses and perception of speech*: 91–101. Fant, G. & Tatham, M.A.A. (Eds). New York: Academic Press.
- Bastian, A. & Schmidt, S. (2008). Affect cues in vocalizations of the bat, *Megaderma lyra*, during agonistic interactions. *J. Acoust. Soc. Am.* **124**, 598–608.

- Beckers, G., Nelson, B. & Suthers, R. (2004). Vocal-tract filtering by lingual articulation in a parrot. *Curr. Biol.* **14**, 1592–1597.
- Beckers, G., Suthers, R. & ten Cate, C. (2003). Pure-tone birdsong by resonance filtering of harmonic overtones. *Proc. Natl. Acad. Sci. USA* **100**, 7372–7376.
- van den Berg, J., Zantema, J. & Doornenbal, P. (1957). On the air resistance and the Bernoulli effect of the human larynx. *J. Acoust. Soc. Am.* **29**, 626–631.
- Blumstein, D. & Armitage, K.B. (1997). Alarm calling in yellow-bellied marmots: I. The meaning of situationally variable alarm calls. *Anim. Behav.* **53**, 143–171.
- Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio in a sampled sound. *Proc. Inst. Phon. Sci. Univ. Amsterdam* **17**, 97–110.
- Bradbury, J.V. & Vehrencamp, S.L. (1998). *Principles of animal communication*. Sunderland: Sinauer.
- Bruckert, L., Liénard, J., Lacroix, A., Kreutzer, M. & Leboucher, G. (2006). Women use voice parameters to assess men's characteristics. *Proc. Roy. Soc. Lond. Ser. B* **273**, 83–89.
- Burdick, C.K. & Miller, J.D. (1975). Speech perception by the chinchilla: discrimination of sustained /a/ and /i/. *J. Acoust. Sci. Am.* **58**, 415–427.
- Caruso, S., Roccasalva, L., Sapienza, G., Zappala, M., Nuciforo, G. & Biondi, S. (2000). Laryngeal cytological aspects in women with surgically induced menopause who were treated with transdermal estrogen replacement therapy. *Fertil. Steril.* **74**, 1073–1079.
- Chan, R.W. & Titze, I.R. (2006). Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics. *J. Acoust. Soc. Am.* **119**, 2351–2362.
- Charlton, B.D. (2008). Female mate choice in nonhuman animals. In *Animal behaviour: new research*: 35–56. Weber, E.A. & Krause, L.H. (Eds). New York: Nova Science Publishers Inc.
- Charlton, B.D., Zhang, Z. & Snyder, R.J. (2009). The information content of giant panda (*Ailuropoda melanoleuca*) bleats: acoustic cues to sex, age and size. *Anim. Behav.* **78**, 893–898.
- Charlton, B.D., McComb, K. & Reby, D. (2008a). Free-ranging red deer hinds show greater attentiveness to roars with formant frequencies typical of young males. *Ethology* **114**, 1023–1031.
- Charlton, B.D., Reby, D. & McComb, K. (2007a). Female perception of size-related formant shifts in red deer, *Cervus elaphus*. *Anim. Behav.* **74**, 707–714.
- Charlton, B.D., Reby, D. & McComb, K. (2007b). Female red deer prefer the roars of larger males. *Biol. Lett.* **3**, 382–385.
- Charlton, B.D., Reby, D. & McComb, K. (2008b). Effect of combined source (F0) and filter (formant) variation on red deer hind responses to male roars. *J. Acoust. Soc. Am.* **123**, 2936–2943.
- Charrier, I., Mathevon, N. & Jouventin, P. (2002). How does a fur seal mother recognize the voice of her pup? An experimental study of *Arctocephalus tropicalis*. *J. Exp. Biol.* **205**, 603–612.
- Charrier, I., Mathevon, N. & Jouventin, P. (2003a). Fur seal mothers memorize subsequent versions of developing pups' calls: adaptation to long-term recognition or evolutionary by-product? *Biol. J. Linn. Soc.* **80**, 305–312.
- Charrier, I., Mathevon, N. & Jouventin, P. (2003b). Vocal signature recognition of mothers by fur seal pups. *Anim. Behav.* **65**, 543–550.
- Cheney, D.L. & Seyfarth, R.M. (1985). Vervet monkey alarm calls: manipulation through shared information? *Behaviour* **94**, 739–751.
- Chuenwattanapranithi, S., Xu, Y., Thipakorn, B. & Manee-wongvatana, S. (2008). Encoding emotions in speech with the size code A perceptual investigation. *Phonetica* **65**, 210–230.
- Clutton-Brock, T. & Albon, S. (1979). The roaring of red deer and the evolution of honest advertising. *Behaviour* **69**, 145–170.
- Cohen, J. & Fox, M. (1976). Vocalizations in wild canids and possible effects of domestication. *Behav. Process* **1**, 77–92.
- Collins, S. (2000). Men's voices and woman's choices. *Anim. Behav.* **60**, 773–780.
- Drahota, A., Costall, A. & Reddy, V. (2008). The vocal communication off different kinds of smile. *Speech Comm.* **50**, 278–287.
- Düppjan, S., Schön, P., Puppe, B., Tuchscherer, A. & Mantuffel, G. (2008). Differential vocal responses to physical and mental stressors in domestic pigs (*Sus scrofa*). *Appl. Anim. Behav. Sci.* **114**, 105–115.
- Elemans, C., Zaccarelli, R. & Herzog, H. (2008). Biomechanics and control of vocalizations in a non-songbird. *J. Roy. Soc. Interface* **5**, 691–703.
- Evans, C.S. (1997). Referential signals. *Perspect. Ethol.* **12**, 99–143.
- Evans, S., Neave, N., Wakelin, D. & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiol. Behav.* **93**, 783–788.
- Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.
- Feighny, J., Williamson, K. & Clarke, J. (2006). North American elk bugle vocalizations: male and female bugle call structure and context. *J. Mammal.* **87**, 1072–1077.
- Fink, B. (1975). *The human larynx: a functional study*. New York: Raven Press.
- Fischer, J., Hammerschmidt, K., Cheney, D. & Seyfarth, R. (2001). Acoustic features of female chacma baboon barks. *Ethology* **107**, 22–54.
- Fischer, J., Hammerschmidt, K., Cheney, D. & Seyfarth, R. (2002). Acoustic features of male baboon loud calls: influences of context, age and individuality. *J. Acoust. Soc. Am.* **111**, 1465–1474.
- Fischer, J., Kitchen, D., Seyfarth, R. & Cheney, D. (2004). Baboon loud calls advertise male quality: acoustic features

- and relation to rank, age, and exhaustion. *Behav. Ecol. Sociobiol.* **56**, 140–148.
- Fisher, D.O., Blomberg, S.P. & Owens, I.P.F. (2002). Convergent maternal care strategies in ungulates and macro-pods. *Evolution* **56**, 167–176.
- Fitch, W.T. (1994). *Vocal tract length perception and the evolution of language*. PhD thesis, Brown University.
- 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. (1999). Acoustic exaggeration of size in birds via tracheal elongation: comparative and theoretical analyses. *J. Zool. (Lond.)* **248**, 31–48.
- Fitch, W.T. (2000a). The evolution of speech: a comparative review. *Trends Cog. Sci.* **4**, 258–267.
- Fitch, W.T. (2000b). The phonetic potential of nonhuman vocal tracts: comparative cineradiographic observations of vocalizing animals. *Phonetica* **57**, 205–218.
- Fitch, W.T. (2000c). Skull dimensions in relation to body size in nonhuman primates: the causal bases for acoustic allometry. *Zoology* **103**, 40–58.
- Fitch, W.T. (2002). Comparative vocal production and the evolution of speech: reinterpreting the descent of the larynx. In *The transition to language*. Wray, A. (Ed.). Oxford: Oxford University Press.
- Fitch, W.T. (2006). Production of vocalizations in mammals. In *Encyclopedia of language and linguistics*: 115–121. Brown, K. (Ed.). Oxford: Elsevier.
- Fitch, W.T. & Fritz, J. (2006). Rhesus macaques spontaneously perceive formants in conspecific vocalizations. *J. Acoust. Soc. Am.* **120**, 2132–2141.
- Fitch, W.T. & 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. & Hauser, M. (2002). Unpacking “honesty”: vertebrate vocal production and the evolution of acoustic signals. In *Acoustic communication*: 65–137. Simmons, A.M., Fay, R.R. & Popper, A.N. (Eds). New York: Springer.
- Fitch, W.T. & Kelley, J. (2000). Perception of vocal tract resonances by whooping cranes *Grus americana*. *Ethology* **106**, 559–574.
- Fitch, W.T., Neubauer, J. & Herzel, H. (2002). Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Anim. Behav.* **63**, 407–418.
- Fitch, W.T. & Reby, D. (2001). The descended larynx is not uniquely human. *Proc. Roy. Soc. Lond. Ser. B* **268**, 1669–1675.
- Fletcher, N., Riede, T., Beckers, G. & Suthers, R. (2004). Vocal tract filtering and the “coo” of doves. *J. Acoust. Soc. Am.* **116**, 3750–3756.
- Fox, M. (1970). A comparative study of the development of facial expressions in canids: wolf, coyote and foxes. *Behaviour* **35**, 49–73.
- Frey, R., Gebler, A., Olson, K.A., Odonkhuu, D., Fritsch, G., Batsaikhan, N. & Stürmer, I.W. (2008). Mobile larynx in Mongolian gazelle: retraction of the larynx during rutting barks in male Mongolian gazelle (*Procapra gutturosa* Pallas, 1777). *J. Morphol.* **269**, 1223–1237.
- Fung, Y. (1981). *Biomechanics: mechanical properties of living tissues*. New York: Springer-Verlag.
- Furrer, R.D. & Manser, M.B. (2009). The evolution of urgency-based and functionally referential alarm calls in ground-dwelling species. *Am. Nat.* **173**, 400–410.
- Ghazanfar, A., Turesson, H., Maier, J., van Dinther, R., Patterson, R. & Logothetis, N. (2007). Vocal tract resonances as indexical cues in the rhesus monkeys. *Curr. Biol.* **17**, 425–430.
- Gouzoules, S., Gouzoules, H. & Marler, P. (1995). Representational signaling in non-human primate vocal communication. In *Current topics in primate vocal communication*: 235–252. Zimmerman, E., Newman, J.D. & Jurgens, U. (Eds). New York: Plenum.
- Halliday, T.R. (1983). The study of mate choice. In *Mate choice*: 3–32. Bateson, P. (Ed.). Cambridge: Cambridge University Press.
- Hardcastle, W. (1976). *Physiology of speech production*. New York: Academic Press.
- Harris, T., Fitch, W., Goldstein, L. & Fashing, P. (2006). Black and white colobus monkey (*Colobus guereza*) roars as a source of both honest and exaggerated information about body mass. *Ethology* **112**, 911–920.
- Hauser, M.D. (1993). The evolution of nonhuman primate vocalizations: effects of phylogeny, body weight and social context. *Am. Nat.* **142**, 528–542.
- Hauser, M.D. (1998). Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Anim. Behav.* **55**, 1647–1658.
- Hauser, M.D. (2000). The sound and the fury: primate vocalizations as reflections of emotion and thought. In *The origins of music*. Wallin, N., Merker, B. & Brown, S. (Eds). Cambridge: MIT Press.
- Hauser, M.D., Evans, C.S. & Marler, P. (1993). The role of articulation in the production of rhesus monkey (*Macaca mulatta*) vocalizations. *Anim. Behav.* **45**, 423–433.
- Hienz, R.D. & Brady, J.V. (1988). The acquisition of vowel discriminations by nonhuman primates. *J. Acoust. Soc. Am.* **84**, 186–194.
- Hirano, M. (1981). *Clinical examinations of voice*. Vienna: Springer Verlag.
- Holekamp, K., Boydston, E., Szykman, M., Graham, I., Nutt, K., Birch, S., Piskiel, A. & Singh, M. (1999). Vocal recognition in the spotted hyaena and its possible implications regarding the evolution of intelligence. *Anim. Behav.* **58**, 383–395.
- Insley, S.J. (2000). Long term vocal recognition in the northern fur seal. *Nature* **406**, 404–405.
- Insley, S.J. (2001). Mother–offspring vocal recognition in northern fur seals is mutual but asymmetrical. *Anim. Behav.* **61**, 129–137.

- Janik, V.M. (2000). Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science* **289**, 1355–1357.
- Janik, V.M., Sayigh, L.S. & Wells, R.S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proc. Natl. Acad. Sci. USA* **103**, 8293–8297.
- Koren, L. & Geffen, E. (2009). Complex call in male rock hyrax (*Procavia capensis*): a multi-information distributing channel. *Behav. Ecol. Sociobiol.* **63**, 581–590.
- Langbein, J. & Putman, R.J. (1992). Reproductive success of female fallow deer in relation to age and condition. In *Biology of deer*: 293–299. Brown, R. (Ed.). New York: Springer-Verlag.
- Le Roux, A., Jackson, T.P. & Cherry, M.I. (2001). The effect of changing call duration and calling bouts on vigilance in Brants' whistling rat, *Parotomys brantsii*. *Behaviour* **138**, 1287–1302.
- Lieberman, P. (1984). *The biology and evolution of language*. Cambridge: Harvard University Press.
- Lieberman, P. (1975). *On the origins of language: an introduction to the evolution of human speech*. New York: Macmillan.
- Lieberman, P. & Blumstein, S.E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.
- Lieberman, P., Klatt, D. & Wilson, W. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkeys and other nonhuman primates. *Science* **164**, 1185–1187.
- Lieberman, P. (1977). *Speech physiology and acoustic phonetics*. Cambridge: Cambridge University Press.
- Macedonia, J.M. (1993). The vocal repertoire of the ringtailed lemur (*Lemur catta*). *Folia Primatol.* **61**, 186–217.
- Macedonia, J.M. & Evans, C.S. (1993). Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology* **93**, 177–197.
- Manser, M. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proc. Roy. Soc. Lond. Ser. B* **268**, 2315–2324.
- Manser, M., Seyfarth, R. & Cheney, D. (2002). Suricate alarm calls signal predator class and urgency. *Trends Cog. Sci.* **6**, 55–57.
- Masataka, N. (1994). Lack of correlation between body size and frequency of vocalisations in young female Japanese macaques (*Macaca fuscata*). *Folia Primatol.* **63**, 115–118.
- Maynard Smith, L. & Harper, D. (2003). *Animal signals*. Oxford: Oxford University Press.
- McComb, K. (1991). Female choice for high roaring rates in red deer, *Cervus elaphus*. *Anim. Behav.* **41**, 79–88.
- McConnell, P. (1990). Acoustic structure and receiver response in domestic dogs, *Canis familiaris*. *Anim. Behav.* **39**, 897–904.
- McConnell, P. & Baylis, J. (1985). Interspecific communication in cooperative herding: acoustic and visual signals from human shepherds and herding dogs. *Z. Tierpsychol.* **67**, 302–382.
- McElligott, A.G., Birrer, M. & Vannoni, E. (2006). Retraction of the mobile descended larynx during groaning enables fallow deer bucks (*Dama dama*) to lower their formant frequencies. *J. Zool. (Lond.)* **270**, 340–345.
- McElligott, A.G. & Hayden, T. (1999). Context-related vocalization rates of fallow bucks, *Dama dama*. *Anim. Behav.* **58**, 1095–1104.
- McElligott, A.G. & Hayden, T. (2001). Postcopulatory vocalisations of fallow bucks: who is listening? *Behav. Ecol.* **12**, 41–46.
- Morton, E. (1977). On the occurrence and significance of motivation-structural rules in some birds and mammal sounds. *Am. Nat.* **111**, 855–869.
- Muller, C. & Manser, M.B. (2008). Mutual recognition of pups and providers in the cooperatively breeding banded mongoose. *Anim. Behav.* **75**, 1683–1692.
- Newton-Fischer, N., Harris, S., White, P. & Jones, G. (1993). Structure and function of red fox *Vulpes vulpes* vocalizations. *Bioacoustics* **5**, 1–31.
- Ohala, J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* **41**, 1–16.
- Ohala, J. (1996). Ethological theory and the expression of emotion in the voice. In *Proc. ICSLP 96, October 3–6, 1996. [4th International Conference on Spoken Language Processing, Philadelphia]*, Vol. 3 pp. 1812–1815. Wilmington: University of Delaware. Available at http://www.linguistics.berkeley.edu/phonlab/users/ohala/papers/emotion_in_the_voice.pdf (accessed 12 June 2008).
- Ohala, J. (2000). *The irrelevance of the lowered larynx in modern Man for the development of speech. Evolution of Language – Paris, ENST (3–6 April 2000)*. Available at http://linguistics.berkeley.edu/phonlab/users/ohala/papers/lowered_larynx.pdf (accessed 5 February 2008).
- Owings, D. & Morton, E. (1998). *Animal vocal communication: a new approach*. Cambridge: Cambridge University Press.
- Owren, J., Seyfarth, R. & Cheney, D. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cyncephalus ursinus*): implications for production processes and functions. *J. Acoust. Soc. Am.* **101**, 2951–2963.
- Owren, M. (1990a). Classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*): I. Natural calls. *J. Comp. Psych.* **104**, 20–28.
- Owren, M. (1990b). Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*): II. Synthetic calls. *J. Comp. Psych.* **104**, 29–40.
- Owren, M. & Bernacki, R. (1998). Applying linear predictive coding (LPC) to frequency-spectrum analysis of animal acoustic signals. In *Animal acoustic communication: Sound analysis and research methods*: 129–161. Hopp, S.L., Owren, M.J. & Evans, C.S. (Eds). New York: Springer-Verlag.
- Palacios, V., Font, E. & Márquez, R. (2007). Iberian wolf howls: acoustic structure, individual variation, and a

- comparison with north American populations. *J. Mammal.* **88**, 606–613.
- Peters, R.H. (1986). *The ecological implications of body size*. Cambridge: Cambridge University Press.
- Pfefferle, D. & Fischer, J. (2006). Sounds and size: identification of acoustic variables that reflect body size in hamadryas baboons, *Papio hamadryas*. *Anim. Behav.* **72**, 43–51.
- Pipitone, R.N. & Gallup, G.G. Jr (2008). Women's voice attractiveness varies across the menstrual cycle. *Evol. Hum. Behav.* **29**, 268–274.
- Pongrácz, P., Molnár, C., Miklósi, A. & Csányi, V. (2005). Human listeners are able to classify dog (*Canis familiaris*) barks recorded in different situations. *J. Comp. Psych.* **119**, 136–144.
- Puts, D.A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evol. Hum. Behav.* **26**, 388–397.
- Puts, D.A., Gaulin, S.J.C. & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evol. Hum. Behav.* **27**, 283–296.
- Puts, D., Hodges, C., Cardenas, R. & Gaulin, S. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evol. Hum. Behav.* **28**, 340–344.
- Reby, D., André-Obrecht, R., Galinier, A. & Farinas, J. (2006). Cepstral coefficients and hidden Markov models reveal idiosyncratic voice characteristics in red deer (*Cervus elaphus*) stags. *J. Acoust. Soc. Am.* **120**, 4080–4089.
- Reby, D., Cargnelutti, B., Joachim, J. & Aulagnier, S. (1999). Spectral acoustic structure of barking in roe deer (*Capreolus capreolus*). Sex-, age- and individual-related variations. *C. R. Acad. Sci. (Life Sci.)* **322**, 271–279.
- Reby, D., Hewison, M., Izquierdo, M. & Pepin, D. (2001). Red deer (*Cervus elaphus*) hinds discriminate between the roars of their current harem-holder stag and those of neighbouring stags. *Ethology* **107**, 951–959.
- Reby, D., Joachim, J., Lek, S., Lauga, J. & Aulagnier, S. (1998). Individuality in the groans of fallow deer (*Dama dama*) bucks. *J. Zool. (Lond.)* **245**, 79–84.
- Reby, D. & 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. & 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. & Clutton-Brock, T. (2005). Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proc. Roy. Soc. Lond. Ser. B* **272**, 941–947.
- Rendall, D. (2003a). Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *J. Acoust. Soc. Am.* **113**, 3390–3402.
- Rendall, D. (2003b). The affective basis of referential grunt variants in baboons. *J. Acoust. Soc. Am.* **113**, 3390–3402.
- 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. *J. Acoust. Soc. Am.* **117**, 944–955.
- Rendall, D., Owren, M. & Rodman, P. (1998). The role of vocal tract filtering in identity cueing in rhesus monkeys (*Macaca mulatta*) vocalizations. *J. Acoust. Soc. Am.* **103**, 602–614.
- Rendall, D., Rodman, P. & Emond, R. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Anim. Behav.* **51**, 1007–1015.
- Rendall, D., Seyfarth, R., Cheney, D. & Owren, M. (1999). The meaning and function of grunt variants in baboons. *Anim. Behav.* **57**, 583–592.
- Rendall, D., Vokey, J. & Nemeth, C. (2007). Lifting the curtain on the Wizard of Oz: biased voice-based impressions of speaker size. *J. Exp. Psychol. Hum. Percept. Perform.* **33**, 1208–1219.
- Riede, T., Arcadi, A.C. & Owren, M.J. (2007). Nonlinear acoustics in the pant hoots of common chimpanzees (*Pan troglodytes*): vocalizing at the edge. *J. Acoust. Soc. Am.* **121**, 1758–1767.
- Riede, T. & Fitch, T. (1999). Vocal tract length and acoustics of vocalisation in the domestic dog (*Canis familiaris*). *J. Exp. Biol.* **202**, 2859–2867.
- Riede, T., Mitchell, B.R., Tokuda, I. & Owren, M.J. (2005). Characterizing noise in nonhuman vocalizations: acoustic analysis and human perception of barks by coyotes and dogs. *J. Acoust. Soc. Am.* **118**, 514–522.
- Riede, T., Owren, M.J. & Arcadi, A.C. (2004). Nonlinear acoustics in pant hoots of common chimpanzees (*Pan troglodytes*): frequency jumps, subharmonics, biphonation, and deterministic chaos. *Am. J. Primatol.* **64**, 277–291.
- Riede, T. & Titze, I.R. (2008). Vocal fold elasticity of the Rocky Mountain elk (*Cervus elaphus nelsoni*) – producing high fundamental frequency vocalization with a very long vocal fold. *J. Exp. Biol.* **211**, 2144–2154.
- Riede, T., Tokuda, I.T., Munger, J.B. & Thomson, S.L. (2008). Mammalian laryngeal air sacs add variability to the vocal tract impedance: physical and computational modeling. *J. Acoust. Soc. Am.* **124**, 634–647.
- Riede, T. & Zuberbühler, K. (2003). The relationship between acoustic structure and semantic information in Diana monkey alarm vocalization. *J. Acoust. Soc. Am.* **114**, 1132–1142.
- Sanvito, S., Galimberti, F. & Miller, E. (2007). Vocal signaling of male southern elephant seals is honest but imprecise. *Anim. Behav.* **73**, 287–299.
- Scherer, K.R. (1986). Vocal affect expression: a review and a model for future research. *Psychol. Bull.* **99**, 143–165.
- Schmidt-Nielsen, K. (1975). Scaling in biology: the consequences of size. *J. Exp. Zool.* **194**, 287–307.
- Schustermann, R.J. & van Parijs, S.M. (2003). Pinniped vocal communication: an introduction. *Aquat. Mamm.* **29**, 177–180.

- Searby, A. & Jouventin, P. (2003). Mother–lamb acoustic recognition in sheep: a frequency coding. *Proc. Biol. Sci.* **270**, 1765–1771.
- Sèbe, F., Aubin, T., Boué, A. & Poindron, P. (2008). Mother–young vocal communication and acoustic recognition promote preferential nursing in sheep. *J. Exp. Biol.* **211**, 3554–3562.
- Seyfarth, R. & Cheney, D. (2003a). Meaning and emotion in animal vocalizations. *Annal. N.Y. Acad. Sci.* **1000**, 32–55.
- Seyfarth, R. & Cheney, D. (2003b). Signalers and receivers in animal communication. *Ann. Rev. Psychol.* **54**, 145–173.
- Seyfarth, R.M. & Cheney, D.L. (1990). The assessment by vervet monkeys of their own and another species' alarm calls. *Anim. Behav.* **40**, 754–764.
- Seyfarth, R.M., Cheney, D.L. & Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* **210**, 801–803.
- Sherman, P.W. (1977). Nepotism and the evolution of alarm calls. *Science* **197**, 1247–1253.
- Singh, S. & Singh, K. (1976). *Phonetics: principles and practices*. Baltimore: University Park Press.
- Sinnott, J.M. (1989). Detection and discrimination of synthetic English vowels by Old World monkeys (*Cercopithecus*, *Macaca*) and humans. *J. Acoust. Soc. Am.* **86**, 557–565.
- Sinnott, J.M. & Kreiter, N.A. (1991). Differential sensitivity to vowel continua in Old World monkeys (*Macaca*) and humans. *J. Acoust. Soc. Am.* **89**, 2421–2429.
- Slocombe, K.E. & Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Curr. Biol.* **15**, 1779–1784.
- Smith, D. & Patterson, R. (2005). The interaction of glottal pulse rate and vocal tract length in judgement of speaker's size, sex and age. *J. Acoust. Soc. Am.* **118**, 3177–3186.
- Solomon, N., Luschi, E. & Liu, K. (1995). Fundamental frequency and tracheal pressure during three types of vocalizations elicited from anesthetized dogs. *J. Voice* **9**, 403–412.
- Sommers, M., Moody, D., Prosen, C. & Stebbins, W. (1992). Formant frequency discrimination by Japanese macaques (*Macaca fuscata*). *J. Acoust. Soc. Am.* **91**, 3499–3510.
- Taylor, A.M., Reby, D. & McComb, K. (2008). Human listeners attend to size information in domestic dog growls. *J. Acoust. Soc. Am.* **123**, 2903–2910.
- Taylor, A.M., Reby, D. & McComb, K. (2009a). Context-related variation in the vocal growling behaviour of the domestic dog (*Canis familiaris*). *Ethol.* **115**, 905–915.
- Taylor, A.M., Reby, D. & McComb, K. (2009b). Size communication in domestic dogs (*Canis familiaris*) growls. *Anim. Behav.*, in press.
- Tembrock, G. (1976). Canid vocalizations. *Behav. Proc.* **1**, 57–75.
- Titze, I.R. (1994). *Principles of vocal production*. Englewood Cliffs: Prentice-Hall.
- Titze, I.R. (2008). Nonlinear source–filter coupling in phonation: theory. *J. Acoust. Soc. Am.* **123**, 2733–2749.
- Tokuda, I., Riede, T., Neubauer, J., Owren, M. & Herzog, H. (2002). Nonlinear analysis of irregular animal vocalizations. *J. Acoust. Soc. Am.* **111**, 2908–2919.
- Torriani, M., Vannoni, E. & McElligott, A.G. (2006). Mother–young recognition in an ungulate hider species: a unidirectional process. *Am. Nat.* **168**, 412–420.
- Trivers, R. (1971). The evolution of reciprocal altruism. *Quart. Rev. Biol.* **46**, 35–57.
- Vannoni, E. & McElligott, A.G. (2007). Individual acoustic variation in fallow deer (*Dama dama*). Common and harsh groans: a source-filter perspective. *Ethology* **113**, 223–234.
- Vannoni, E. & McElligott, A.G. (2008). Low frequency groans indicate larger and more dominant fallow deer (*Dama dama*) males. *PLoS ONE* **3**, e3113.
- Vannoni, E., Torriani, M. & McElligott, A.G. (2005). Acoustic signaling in cervids: a methodological approach for measuring vocal communication in fallow deer. *Cog. Br. Behav.* **IX**, 551–565.
- Volodin, I.A., Volodina, E.V. & Efremova, K.O. (2009). Antelope, calling through the nose: structure of sounds and effect of sexual selection on the vocal behavior of the saiga (*Saiga tatarica*). *Zool. Zhurnal* **88**, 113–124.
- Volodina, E., Volodin, I., Isaeva, I. & Unck, C. (2006). Biphonation may function to enhance individual recognition in the dhole, *Cuon alpinus*. *Ethology* **112**, 815–825.
- Wilden, I., Herzog, H., Peters, G. & Tembrock, G. (1998). Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics* **9**, 171–196.
- Wyman, M.T., Mooring, M.S., McCowan, B., Penedo, M.C.T. & Hart, L.A. (2008). Amplitude of bison bellows reflects male quality, physical condition and motivation. *Anim. Behav.* **76**, 1625–1639.
- Yin, S. (2002). A new perspective on barking in dogs (*Canis familiaris*). *J. Comp. Psychol.* **116**, 189–193.
- Yin, S. & McCowan, B. (2004). Barking in domestic dogs: context specificity and individual identification. *Anim. Behav.* **68**, 343–355.
- Zimmermann, E. & Lerch, C. (1993). The complex acoustic design of an advertisement call in male-mouse lemurs (*Microcebus murinus*, *prosimii*, primates) and sources of its variation. *Ethology* **93**, 211–224.
- Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. *Anim. Behav.* **63**, 293–299.