

REVIEW

Vocal expression of emotions in mammals: mechanisms of production and evidence

E. F. Briefer

Biological and Experimental Psychology Group, School of Biological and Chemical Sciences, Queen Mary University of London, London, UK

Keywords

acoustic communication; affective state; arousal; motivation; source-filter theory; stress; valence; vocal production.

Correspondence

Elodie F. Briefer, Biological and Experimental Psychology Group, School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, UK.

Email: e.briefer@qmul.ac.uk

Editor: Steven Le Comber

Received 23 January 2012; revised 18 March 2012; accepted 29 March 2012

doi:10.1111/j.1469-7998.2012.00920.x

Abstract

Emotions play a crucial role in an animal's life because they facilitate responses to external or internal events of significance for the organism. In social species, one of the main functions of emotional expression is to regulate social interactions. There has recently been a surge of interest in animal emotions in several disciplines, ranging from neuroscience to evolutionary zoology. Because measurements of subjective emotional experiences are not possible in animals, researchers use neurophysiological, behavioural and cognitive indicators. However, good indicators, particularly of positive emotions, are still lacking. Vocalizations are linked to the inner state of the caller. The emotional state of the caller causes changes in the muscular tension and action of its vocal apparatus, which in turn, impacts on vocal parameters of vocalizations. By considering the mode of production of vocalizations, we can understand and predict how vocal parameters should change according to the arousal (intensity) or valence (positive/negative) of emotional states. In this paper, I review the existing literature on vocal correlates of emotions in mammals. Non-human mammals could serve as ideal models to study vocal expression of emotions, because, contrary to human speech, animal vocalizations are assumed to be largely free of control and therefore direct expressions of underlying emotions. Furthermore, a comparative approach between humans and other animals would give us a better understanding of how emotion expression evolved. Additionally, these non-invasive indicators could serve various disciplines that require animal emotions to be clearly identified, including psychopharmacology and animal welfare science.

Introduction

The existence of emotions in animals had already been suggested by Darwin in his book *The Expression of the Emotions in Man and Animals* (Darwin, 1872). An emotion is not a high-level cognitive process, as evidence suggests that emotional states are in fact generated by lower (medial and caudal subcortical structures) rather than higher brain regions (neocortical structures; Panksepp, 2005). Emotions have a crucial function for an animal's life as they facilitate responses to external or internal events of significance for the organism; positive emotions elicit approach behaviour towards stimuli that enhance fitness ('rewards'), whereas negative emotions trigger avoidance behaviour when encountering stimuli that threaten fitness ('punishers'; Mendl, Burman & Paul, 2010).

In scientific terms, an emotion is an intense but short-living affective reaction to a specific event or stimulus. However, for most people, 'emotion' is a synonym of 'feeling' (i.e. our conscious/subjective experience of emotions). For example, if we happen to encounter a dangerous animal in the wild, our

heart rate will increase and we will begin to sweat. Our subjective feeling of these physiological changes is what we call 'fear' (Davidson, Scherer & Goldsmith, 2003). This is probably why I often hear people asking 'do animals really have emotions?' or 'is it not being anthropomorphic to infer that animals have emotions?' Yes, non-human animals have emotions (at least 'basic emotional systems': seeking, rage, fear, lust, care, panic and play; Panksepp, 2011), even if subjective emotional experiences are not yet possible to prove in animals (de Waal, 2011). In other words, animals express signs of emotions, but their ability to feel these emotions is still highly controversial (Panksepp, 2005).

Studying animal emotions can reveal the nature of basic human emotions (Panksepp, 2011). It can help us to understand how emotions evolved and developed, in order to acquire a full understanding of their nature (Adolphs, 2010). Knowing which emotions animals are experiencing could also serve several disciplines such as evolutionary zoology, affective neuroscience, pharmaceuticals (pain research) and comparative psychology. Today, public concern about animal

welfare is strongly based on the attribution of mental states to animals, and welfare assessment is now commonly linked to both physical and mental health (Dawkins, 2008). The problem then is how can we measure emotions in animals if they cannot tell us what they feel (i.e. subjective component)? A robust framework to study animal emotional states has recently been established by Mendl *et al.* (2010). This framework suggests using the other components of emotion as indicators; neurophysiological, behavioural and cognitive components, and the two dimensions of emotions; arousal (i.e. intensity or activating qualities) and valence (i.e. positivity/negativity). Therefore, now, animal research is on the right path towards a full understanding of animal emotions. However, the proposed neurophysiological, behavioural and cognitive indicators of emotions need to be described in detail before we are able to infer animal emotions.

Facial expressions of emotions have been studied in several animal species (e.g. non-human primates, sheep *Ovis aries*, rats *Rattus norvegicus*; Tate *et al.*, 2006; Langford *et al.*, 2010). Another promising behavioural indicator of emotions is vocalizations. Several types of vocalizations have been shown to indicate positive or negative emotional valence (e.g. ultrasonic vocalizations in rats; Knutson, Burgdorf & Panksepp, 2002; Burgdorf, Panksepp & Moskal, 2011). Their link to specific brain circuits responsible for emotions has been established in some species (e.g. cats *Felis catus*, Siegel *et al.* 2010; rats, Burgdorf *et al.*, 2007). However, the link between variations in vocal parameters and emotion-related physiological changes in the vocal apparatus has rarely been investigated. In humans, indicators of emotions in human voice ('affective prosody') have been studied in detail (e.g. Scherer, 1986; Zei Pollermann & Archinard, 2002). Theories of speech production recently applied to animal vocal communication ('source-filter theory of vocal production'; Fant, 1960; Titze, 1994; Taylor & Reby, 2010) can inform us about the mechanisms linking contexts of vocal production and the acoustic structure of vocalizations, and allow us to make predictions about how vocalizations should change according to emotional arousal and valence.

In this paper, I review the current state of knowledge on vocal correlates of emotions in mammals. I first introduce techniques recently developed to study animal emotions. Then, I describe methods used to study animal vocalizations, which link vocal parameters to production mechanisms. In the following sections, I review the existing literature on vocal correlates of emotions in humans and other mammals. I highlight the best methods to use in studies on non-human mammals, and the lack of research in this area. Finally, I conclude with the best/most likely vocal indicators of emotional valence and arousal in non-human mammals.

Measuring emotions in animals

Measuring animal emotions might appear, at first glance, as a difficult goal to achieve. Fortunately, the interest in the field of affective biology has considerably increased recently. As a result, new frameworks have emerged, offering researchers convenient and accurate techniques to measure animal

emotional states, including positive emotions and moods (i.e. long-term diffuse emotional states that are not directly caused by an event; e.g. Désiré, Boissy & Veissier, 2002; Paul, Harding & Mendl, 2005; Boissy *et al.*, 2007; Mendl *et al.*, 2010). The basic principle behind those measures is relatively simple: an animal is assumed to experience a given emotion (e.g. fear) if it shows neurophysiological (e.g. changes in brain activity or in heart rate), behavioural (e.g. facial expression, production of calls, fleeing behaviour) and/or cognitive (e.g. increase in attention towards the stimulus, 'attention bias') signs of this emotion in a situation presumed to induce it. Therefore, to study a given emotion, a first step consists in placing the animal in a situation presumed to trigger this emotion and then measuring the corresponding pattern of neurophysiological, behavioural and/or cognitive changes induced. The resulting emotion-specific profile of responses can then be used later as evidence that the emotion is elicited in other situations.

I will present here the framework developed by Mendl *et al.* (2010), one of several useful theories within this field to study emotions (e.g. see also appraisal theories; Désiré *et al.*, 2002). This framework proposes to assess emotions using the measurable components of the organism's emotional response (neurophysiological, behavioural and cognitive) through the two dimensions of emotions (valence and arousal; 'dimensional approach'). As opposed to the 'discrete emotion approach', which suggests the existence of a small number of fundamental emotions associated with very specific neurophysiological response patterns, the 'dimensional approach' suggests that all types of emotions can be mapped in the space defined by valence and arousal (i.e. by a given combination of these two dimensions). Therefore, neurophysiological, behavioural and cognitive measures reliably associated with a particular location in this two-dimensional space can be used as indicators of the emotion defined by this location. For example, indicators of 'fear' will be components reliably associated with negative valence and high arousal, whereas those of 'contentment' will be components reliably associated with positive valence and low arousal (Mendl *et al.*, 2010). This approach is useful for the study of animal emotions because it allows researchers to investigate differences between emotional states of low versus high arousal and of positive versus negative valence, without having to infer the specific emotion that the animal is experiencing.

Commonly used neurophysiological indicators are stress measures, such as heart rate and its variability, respiration rate, skin temperature, electrodermal response or neuroendocrine activity. Good behavioural indicators include body postures, movements and vocalization types and rate (e.g. Reefmann *et al.*, 2009a; Reefmann, Wechsler & Gyax, 2009b). Other related techniques allow researchers to assess animal long-term emotional states ('moods') using the cognitive components of emotions, such as appraisal processes and attention, memory and judgment biases (Paul *et al.*, 2005). The studies carried out so far show that it might be difficult to differentiate between situations of similar arousal, but different valence (Mendl *et al.*, 2010). Considering multiple indicators could help to interpret emotions experienced by animals

(Paul *et al.*, 2005; Boissy *et al.*, 2007). Therefore, new indicators are needed, especially to distinguish between positive and negative emotional valence.

Measuring vocalizations in mammals

Research on mammal vocal communication, and particularly studies on vocal indicators of emotions and welfare, often focused principally on the most obvious parameters of vocalizations, such as calling rate, duration, the occurrence of call types and energy distribution (e.g. Weary & Fraser, 1995a; Weary, Braithwaite & Fraser, 1998; Byrne & Suomi, 1999; Grandin, 2001; Marchant, Whittaker & Broom, 2001; Shair *et al.*, 2003). The types of vocalizations produced can be useful indicators of emotional arousal and valence (Brudzynski, 2007; Scheumann, Zimmermann & Deichsel, 2007; Taylor, Reby & McComb, 2009; Gogoleva *et al.*, 2010a). However, new methods, adapted from studies on human speech to non-human mammal vocalizations, could allow a far better understanding of why and to what extent calls vary between individuals and between contexts (Taylor & Reby, 2010).

According to the source-filter theory of voice production (Fant, 1960; Titze, 1994), mammal vocalizations are generated by vibrations of the vocal folds ('source') and are subsequently filtered in the vocal tract ('filter'). The source determines the fundamental frequency of the call (*F0*; vocal measures mentioned throughout the review are in *italic* and their definitions are listed in Table 1), and the filter shapes the source signal by selectively amplifying certain frequencies and dampening out others. This filtering mechanism produces spectral peaks called '*formants*' (Fig. 1). Source-related vocal parameters depend on the anatomy and physiology of the larynx (vocal fold length, thickness, mass, tension and internal structure, i.e. collagen and elastin fibre densities and orientations), whereas filter-related vocal parameters are determined by the anatomy and physiology of the supralaryngeal vocal tract (e.g. shape, length and tension of the vocal tract; Table 2). The source-filter theory has recently been applied to various species and revealed interesting links between vocalizations and the caller's anatomical or physiological attributes (e.g. Reby & McComb, 2003; Briefer, Vannoni & McElligott, 2010; Briefer & McElligott, 2011; Charlton *et al.*, 2011).

The source-filter framework could help in predicting and identifying parameters influenced by emotions because it considers the link between the structure of vocalizations and their mode of production. In animals as in humans, very few studies on emotions have investigated the frequency distribution in the spectrum or formant parameters (Scherer, 2003; Juslin & Scherer, 2005). However, several studies have suggested that this could be key to the vocal differentiation of emotional valence, with the other parameters (e.g. *F0*, *amplitude* and *vocalization rate*) indicating mainly physiological arousal (Scherer, 1986; Banse & Scherer, 1996; Waaramaa *et al.*, 2010; Patel *et al.*, 2011). Therefore, it is crucial to measure a large set of parameters including *formant frequencies*, using the source-filter framework, in order to obtain emotion-specific vocal profiles. In the next sections, I will review the literature on vocal correlates of emotions in humans and other mammals,

and explain how both *F0 contour* and *formants* can be influenced by the emotional state of the caller.

Vocal correlates of emotions in humans

Human speech communicates both linguistic and paralinguistic (i.e. non-verbal; voice quality and prosody) information. Because only equivalents of non-verbal cues can be found in non-human mammals, I focus in this review on emotion indicators in the paralinguistic domain. In humans, vocal correlates of emotions in this domain ('affective prosody') play an important role in social interactions, and have been extensively studied since Darwin (1872). Both the encoding (expression) and the decoding (impression) of discrete emotions in the voice have been studied (Banse & Scherer, 1996). Research on the coding process has revealed a set of acoustic characteristics that reliably indicate emotions (see next sections for more details; Zei Pollermann & Archinard, 2002; Scherer, 2003). The specific acoustic profile of several different emotions, showing similarities across languages, has been established (Hammerschmidt & Jürgens, 2007; Pell *et al.*, 2008). Studies on the decoding process have shown that people are able to extract accurate information about discrete emotions from vocal cues, even across cultures and languages (Scherer, Banse & Wallbott, 2001; Sauter *et al.*, 2010).

How is speech produced?

Speech is produced through the processes of respiration, phonation, resonance and articulation (see Table 2; Fant, 1960; Titze, 1994; Juslin & Scherer, 2005). The lungs generate an air flow, which then passes through the larynx. In the larynx, the air flow is converted into sound by vibration of the vocal folds. Then, this sound is filtered in the supralaryngeal vocal tract (pharynx, oral and nasal cavities), before radiating into the environment through the lips and nostrils. We therefore have three systems involved in the production of speech. The respiratory system (respiration process) includes the lungs and determines the *duration*, *rate*, *amplitude*, and the subglottal pressure, which influences *F0*. The phonatory system (phonation process) includes the larynx and all sub-laryngeal and laryngeal structures. This system determines the characteristics of the source signal (*F0 contour*; 75–300 Hz for men, 100–500 Hz for women). Finally, the filter system (resonance and articulation processes) includes all the air cavities between the larynx and the opening of the mouth and nostrils (vocal tract) and determines the *energy distribution* of the sound (*frequency spectrum* characteristics and *formant contour*). The structure of vocalizations therefore depends on the anatomy and physiology of each of these systems.

The mechanism of vocal production is similar in humans and other mammals. However, in humans, the particular position of the larynx that rests low in the throat and is also mobile, gives us a long and flexible pharyngeal cavity and a nearly 90° connection between the pharyngeal and oral

Table 1 Vocal parameters listed in this review and their description

Code	Parameter	Description
1	Vocalization/element duration	Total duration of the vocalization/elements in the vocalization
2	Vocalization/element rate	Number of vocalizations/elements produced per time unit
3	Inter-vocalization/element interval	Mean silence duration between two vocalizations/elements
4	Number of elements	Number of syllable/note composing the vocalization (for complex vocalizations)
5	<i>F0</i> <i>F0 contour</i>	Fundamental frequency, lowest frequency of the vocalization Sequence of <i>F0</i> values across the vocalization (includes <i>F0</i> mean, start, end, minimum, maximum)
6	<i>F0</i> range	Difference between minimum and maximum <i>F0</i>
7	Time of maximum/minimum <i>F0</i>	Time point of the maximum/minimum <i>F0</i> value relative to the total duration
8	H1-H2	Difference in amplitude level between <i>F0</i> and the second harmonic ('hoarseness' or breathiness' in human voice)
9	<i>F0</i> slope	<i>F0</i> mean absolute slope (steepness)
10	Jitter	Cycle-to-cycle frequency variation of <i>F0</i>
11	Shimmer	Cycle-to-cycle amplitude variation of <i>F0</i>
12	Amplitude Amplitude contour	Level of energy in the vocalization (intensity or energy) Sequence of amplitude values across the vocalization (includes mean, start, end, minimum, maximum amplitude)
13	Amplitude range	Difference between minimum and maximum amplitude
14	Amplitude modulation (AM)	Variation in amplitude relative to the total duration
15	Frequency spectrum Energy distribution	Amplitude as a function of frequency Distribution of energy in the spectrum (e.g. energy quartiles, amount of energy in various part of the spectrum, ratio between harmonics and <i>F0</i>)
16	Peak frequency	Frequency of maximum amplitude
17	Time of peak frequency	Time point of the frequency peak relative to the total duration
18	Dominant frequency band (F _{Bn}) contour	Sequences of values of high amplitude frequency bands (F _{B1} , 2, 3, etc) across the vocalization (e.g. F _{Bn} mean, start, end, minimum, maximum)
19	Spectral slope	Slope of the regression line through the spectrum
20	Frequency range	Frequency range in the spectrum (e.g. difference between energy quartiles)
21	Frequency modulation (FM) <i>Formants</i> frequencies <i>F1</i> , 2, 3, 4, etc	Variability/modulation of the dominant frequency or <i>F0</i> across the call Concentration of acoustic energy around particular frequencies in the vocalization Frequency value of the first, second, third, fourth, etc. formant
22	<i>F_n</i> contour	Sequences of values of <i>formant frequencies</i> (<i>F1</i> , 2, 3, etc) across the vocalization (e.g. <i>F_n</i> mean, start, end, minimum, maximum)
23	Formant dispersion Non-linear phenomena	Spacing of the <i>formants</i> Complex intrusions into the normal spectral structure (e.g. subharmonics, deterministic chaos, biphonation, frequency jumps)
24	Spectral noise	Proportion of noise in the vocalization, where the harmonic structure is not clear or cannot be detected (e.g. chaos)
25	Entropy	Ratio of the geometric mean to the arithmetic mean of the spectrum (0: pur tone; 1: random noise)
26	Harmonic-to-noise ratio	Ratio of amplitude peaks of detectable harmonics to noise threshold (higher values indicate more tonal vocalizations)
27	Subharmonics	Proportion of the total duration with additional spectral components in the harmonic series (fractional multiples of <i>F0</i> , e.g. <i>F0</i> /2, <i>F0</i> /3, 2/3 <i>F0</i> , 4/3 <i>F0</i> , etc.)

These parameters correspond to the most commonly measured parameters in studies on non-human mammals. Different names were given for the same parameters across studies, and same names were sometimes attributed to different parameters. I chose the names listed here because they are the most commonly attributed to each parameter listed. Names are kept constant throughout the review, and might differ from the names given in the references cited, but correspond closely to the description given by the authors.

cavities. Consequently, we benefit from important articulatory possibilities. We are able to modify the size of our oral and pharyngeal cavity using our tongue, lips, teeth, hard and soft palate, and jaw. This ability plays a crucial role in human speech. For example, by constricting the vocal tract in different places, we can create various patterns of change in the first

two *formants* (*F1*, around 500 Hz; *F2*, around 1500 Hz), thus producing different vowels. Higher *formants* (e.g. *F3*, around 2500 Hz) are fairly constant and depend on the vocal tract length (Fant, 1960). These morphological particularities associated with an important motor control are at the basis of the evolution of speech (Fitch, 2000a; Jürgens, 2009).

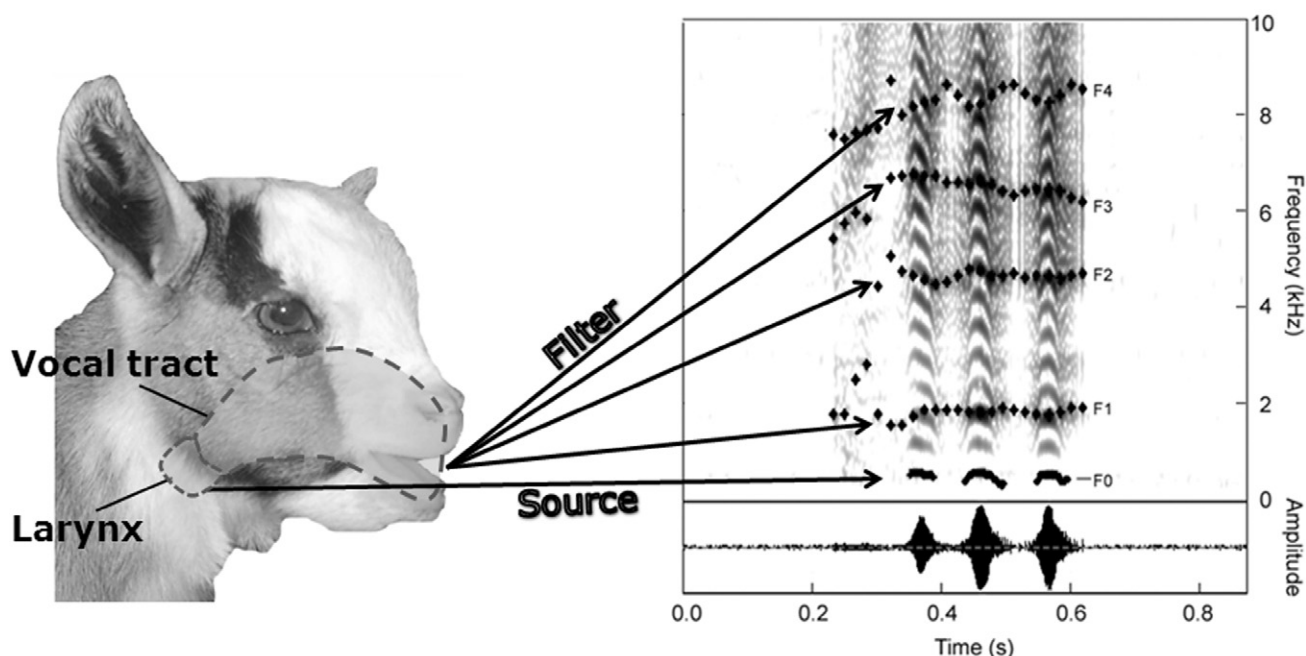


Figure 1 Illustration of the source–filter theory of vocal production. Left: schema of the vocal production mechanism in a goat kid indicating the approximate location of the larynx and vocal tract (including nasal cavities). Right: spectrogram (above) and oscillogram (below) of the corresponding call showing the fundamental frequency (F_0 , black line at the bottom of the spectrogram) and the first four *formants* (F_1 – F_4 , black dots above). The source sound is produced in the larynx by vibration of the vocal folds, determining the fundamental frequency of the call (F_0) and the harmonics. This source sound is then filtered in the vocal tract, determining the *formants* (F_1 – F_4), which correspond to a concentration of acoustic energy around particular frequencies. The positions of the larynx and of vocal tract have been estimated following (Fitch, 2000b).

Table 2 Vocal production mechanism in mammals

	System	Location	Function in vocal production	Associated vocal parameters
Source	Respiration	Lungs and trachea	Generating and conducting the air flow	Amplitude, duration, F_0 (subglottal pressure)
	Phonation	Larynx (including vocal folds)	Transforming the air flow into sound by oscillation of the vocal folds	F_0
Filter	Resonance	Vocal tract (pharynx, vocal, nasal and oral cavities, lips and nostrils)	Filtering the source sound by amplifying and attenuating certain frequencies	<i>Formants</i> , relative energy distribution in the spectrum
	Articulation (humans)	Tongue, lips, hard and soft palate, teeth and jaw	Transforming the incoming sound in language-specific speech sounds (unvoiced/voiced sounds)	F_1 and F_2 contours, relative energy distribution in the spectrum

Systems of sound production and corresponding anatomy (location), function in vocal production and vocal parameters associated.

How is affective prosody studied?

Three types of research paradigms have been used to study affective prosody in humans: natural vocal expression, induced emotional expression and simulated emotional expression (Murray & Arnott, 1993; Scherer, 2003; Juslin & Scherer, 2005). The first approach consists of analysing voices recorded in naturally occurring emotional situations and is of high ‘ecological validity’ (i.e. high accuracy of the underlying speaker state; e.g. Williams & Stevens, 1972; Roessler &

Lester, 1976; Frolov *et al.*, 1999). The second approach is based on artificially induced emotions in the laboratory, using psychoactive drugs, presentation of emotion-inducing films or images, or recall of emotional experiences (e.g. Scherer *et al.*, 1985; Tolkmitt & Scherer, 1986; Zei Pollermann & Archinard, 2002). The third and most often used approach consists of analysing simulated emotional expression, produced by actors asked to pronounce a word or sentence by expressing particular emotional states (e.g. van Bezooijen, 1984; Banse & Scherer, 1996; Hammerschmidt & Jürgens, 2007).

How do emotions influence speech parameters?

Vocal cues to emotions are emitted involuntarily. To summarize, emotions induce changes in the somatic and autonomic nervous system (SNS and ANS), which in turn cause tension and action of muscles used for voice production (phonation, resonance and articulation), as well as changes in respiration and salivation. All these modifications to the vocal apparatus result in particular changes of voice parameters (Scherer, 2003). The SNS is more directly involved in motor expression, whereas the ANS mainly impacts on respiration and the secretion of mucus and salivation (Scherer, 1986). The impacts of the ANS on vocalizations will depend on the respective dominance of the sympathetic (ergotropic) and parasympathetic (trophotropic) branches, which differs between emotions (Zei Pollermann, 2008). High-arousal emotions are associated with a high sympathetic tone and a low parasympathetic tone, and the opposite applies to low-arousal emotions.

A change in respiration can cause changes in speech *duration*, *amplitude* and *rate*, as well as in *F0* by increasing the subglottal pressure (i.e. pressure generated by the lungs beneath the larynx). An increase in the action and/or tension of the respiratory muscles can induce longer *durations*, higher *amplitude* and higher *F0*. Salivation acts on the resonance characteristics of the vocal tract, with a decrease in salivation resulting in higher *formant frequencies* (Scherer, 1986; Zei Pollermann & Archinard, 2002). The effects of the main muscles are as follows. In the larynx, an increase in the action and/or tension of the cricothyroid muscles stretches the vocal folds, resulting in higher *F0*, whereas an increase in action and/or tension of the thyroarytenoid muscles shorten and thicken the vocal folds, resulting in a lower *F0* (Titze, 1994). The actions of the sternothyroid and sternohyoid muscles pull the larynx downward, resulting in an elongation of the vocal tract length, and therefore lower *formant frequencies*. Pharyngeal constriction, and tension of the vocal tract walls, result in an increase of the proportion of energy in the upper part of the *frequency spectrum* (above 500 Hz) in relation to the energy in the lower frequency region, i.e. a shift in *energy distribution* towards higher frequencies. By contrast, pharyngeal relaxation results in an increase of the proportion of energy in the lower part of the *frequency spectrum* (below 500 Hz; Scherer, 1986). The relative raising or lowering of the *formants* (*F1*, *F2*, *F3*, etc.) depends on the length of the vocal tract, the configuration of the pharyngeal regions and oral and nasal cavities, and the opening of the mouth. Increased mouth opening raises *F1* closer to *F2*. In the case of pharyngeal constriction and mouth retraction, *F1* should rise and *F2* and *F3* should fall. Finally, protrusion of the lips increases the length of the vocal tract, lowering all *formant frequencies* (Fant, 1960; Fitch & Hauser, 1995).

Physiological arousal is mainly reflected in parameters linked to respiration and phonation, such as *F0*, *amplitude* and timing parameters (e.g. *duration* and *rate*), while emotional valence seems to be reflected in intonation patterns and voice

quality (i.e. pattern of energy distribution in the spectrum; Scherer, 1986). Emotions of high arousal, such as fear or joy, are associated with an increase in *amplitude*, *F0*, *F0 range*, *F0 variability*, *jitter*, *shimmer* and *speech rate*, as well as with fewer and shorter interruptions (*inter-vocalization interval*). By contrast, emotions of low arousal, such as boredom, induce a low *F0*, narrow *F0 range* and low *speech rate* (Scherer, 1986; Murray & Arnott, 1993; Bachorowski & Owren, 1995; Banse & Scherer, 1996; Zei Pollermann & Archinard, 2002; Juslin & Scherer, 2005; Li *et al.*, 2007). A recent study showed that source-related parameters linked to phonatory effort (tension), perturbation and voicing frequency allowed good classification of five emotions (relief, joy, panic/fear, hot anger and sadness), but did not allow good differentiation of emotional valence (Patel *et al.*, 2011). Filter-related cues (*energy distribution*, *formant frequencies*) have been more rarely considered in studies of emotions (Juslin & Scherer, 2005). However, it seems that spectrum parameters, particularly the *energy distribution* in the spectrum, *F3* and *F4*, contrary to source-related parameters, differ between emotions of similar arousal but different valence (Banse & Scherer, 1996; Laukkanen *et al.*, 1997; Zei Pollermann & Archinard, 2002; Waaramaa, Alku & Laukkanen, 2006; Waaramaa *et al.*, 2010). Emotion perception studies showed that an increase in *F3* is judged as more positive (Waaramaa *et al.*, 2006, 2010). Valence could also be reflected in other voice quality- and amplitude-related parameters, with positive emotions being characterized by steeper *spectral slopes*, narrower *frequency ranges*, less noisy signals (*spectral noise*), lower *amplitude* levels and earlier positions of the *maximum peak frequency* than negative ones (Hammerschmidt & Jürgens, 2007; Goudbeek & Scherer, 2010). Furthermore, the energy is lower in frequency in positive compared with negative emotions in a large portion of the spectrum (Zei Pollermann & Archinard, 2002; Goudbeek & Scherer, 2010).

Difficulties of studying vocal correlates of emotion in humans

There are several difficulties associated with the study of affective prosody in humans. First, voice parameters do not only result from the physiological state of the speaker, but also from socio-cultural and linguistic conventions, and more generally from voluntary control of emotion expression. Therefore, psychological, social interactional and cultural determinants of voice production may counteract each other, and act as confounding factors in the study of affective prosody (Scherer, Ladd & Silverman, 1984; Scheiner & Fisher, 2011). Second, interferences can exist between linguistic and paralinguistic domains (i.e. between vocal emotion expression and semantic or syntactic cues). In particular, the investigation of the role of *formants* in emotional communication is rendered difficult by their linguistic importance. They have been suggested to be crucial for communicating emotional valence, but this hypothesis is difficult to test in humans (Laukkanen *et al.*, 1997; Waaramaa *et al.*, 2010). Third, it is very difficult to study emotional processes in natural situations or to experimentally induce

strong emotional states in the laboratory (Scherer, 1986). Finally, another problem is that most studies have investigated correlates of discrete emotions ('discrete emotion approach', as opposed to the 'dimensional approach'), despite a lack of qualitative description of basic emotions. Emotion terms are rather imprecise, do not systematically correspond to emotional states and differ between languages, which renders the overall description of vocal expression of emotion complex (Scherer, 1986; Murray & Arnott, 1993). Most of these problems might not be present in non-human animals, in which vocalizations are supposed to be under lower voluntary control than in human. Animal vocalizations should reflect emotions more directly, free of conventionalization or self-presentation constraints (Jürgens, 2009). Therefore, vocal correlates of emotions in animals could serve as an interesting, simplified model of human affective prosody and provide evidence of a phylogenetic continuity of emotion vocalizations (Scherer, 2003; Juslin & Scherer, 2005).

Vocal correlates of emotions in non-human mammals

In animals as in humans, cues to emotional states (e.g. visual, vocal) regulate social interactions, because they inform individuals about the probable intentions of behaviours of others (Panksepp, 2009; Keltner & Lerner, 2010). Therefore, vocal correlates of emotions have a crucial function in social species (Brudzynski, 2007). Vocal production mechanisms being very similar between humans and other mammals, comparable changes in vocal parameters in response to emotional states are expected (Scherer & Kappas, 1988; Manteuffel, Puppe & Schön, 2004; Scheiner & Fisher, 2011). Unlike the research on humans described earlier, there has been a lack of studies on the effects of emotions on vocalizations in other mammals, despite these effects being mentioned already by Darwin (1872). By contrast, the effect of motivation on animal vocalizations has been widely studied, since the concept of 'motivation-structural rules' described by Collias (1960) and Morton (1977). According to this concept, vocalizations produced in 'hostile' contexts should be structurally different from those produced in 'friendly' or 'fearful' contexts (Morton, 1977). Motivation states differ from emotions in the sense that they refer to the likelihood that an animal would perform a certain behaviour (e.g. attack, retreat), and not directly to its emotional state (Zahavi, 1982). Vocal correlates of motivation can be defined as 'strategic use of expressive displays independent of the presence of appropriate internal determinants, based on ritualized meanings of state-display relations' (Scherer, 1986). Nevertheless, they imply an underlying emotion. For example, a call emitted in a 'friendly' context implies that the producer of the call is in a positive emotional state. Therefore, findings related to motivation-structural rules can be used to predict how vocal parameters should vary according to emotions. In the next part, I describe the concept of motivation-structural rules and find-

ings in this area of research, before reviewing the literature on vocal correlates of emotions.

Motivation-structural rules

Motivation-structural rules emerged from the comparison between vocalizations produced by numerous species of birds and mammals. Morton (1977) observed that the acoustic structure of calls can often be predicted from the context of production. In hostile contexts, animals generally produce low-frequency calls. Morton suggested that because low-frequency calls mimic large-sized animals, their production increases the perceived size of the caller during hostile interactions. By contrast, high tonal sounds are produced in fearful or appeasing contexts. Because they mimic the sounds produced by infants, these sounds should have an appeasing effect on the receiver(s). Accordingly, intermediate stages between hostility and fear or appeasement are characterized by intermediate call frequencies. Since Morton (1977), this hypothesis has been tested in several species [e.g. African wild dog *Lycan pictus* (Robbins & McCreery, 2003) chimpanzee *Pan troglodytes* (Siebert & Parr, 2003) coati *Nasua narica* (Compton *et al.*, 2001) dog *Canis familiaris* (Yin & McCowan, 2004; Pongrácz, Csaba & Miklósi, 2006; Lord, Feinstein & Coppinger, 2009; Taylor *et al.*, 2009) grey mouse lemur *Microcebus murinus* (Scheumann *et al.*, 2007) North American elk *Cervus elaphus* (Feighny, Williamson & Clarke, 2006) white-faced capuchins *Cebus capucinus* (Gros-Louis *et al.*, 2008) white-nosed macaques *Macaca spp.* (Gouzoules & Gouzoules, 2000)]. Most of these studies showed that, in accordance with the motivation-structural rules, calls produced during agonistic encounters are of long durations, with low frequencies, wide frequency ranges and little frequency modulations. Conversely, calls produced during non-aggressive behaviour, or fearful situations, are often of short durations, tonals (no spectral noise), with high frequencies and frequency modulations. Therefore, call structure can be partially predicted by the motivation-structural rules in numerous species (August & Anderson, 1987).

The variation between motivational call types could reflect different emotional valences, whereas the variation within motivational call types is probably due to differences in arousal states (Manser, 2010). If we logically assume that an individual in a hostile context is experiencing a negative emotional state of high arousal, whereas an individual in a friendly context is experiencing a positive emotional state of high arousal, then negative emotions could be characterized by low-frequency sounds and positive emotions by high-frequency sounds. However, the theory predicts that high-frequency sounds are also produced in fearful contexts, which assume a negative emotional state of high arousal. According to August & Anderson (1987), fearful and friendly contexts represent two very different motivation states, and could be distinguished by measuring more acoustic parameters than those suggested by Morton (1977). The relationship between emotions and call structure might not be entirely predicted from the motivation-structural rules, but the opposite could be true (i.e. motivation-structural rules could be explained by

the underlying emotional state of the caller in aggressive/friendly contexts). Therefore, vocal correlates of emotions need to be studied using experimental situations, specifically designed to trigger emotions characterized by a given valence and arousal.

Evidence of vocal expression of emotion

I carried out an extensive search of the available literature with the following keywords: vocal, expression, communication, call, acoustic, mammal, animal, condition, context, stress, welfare, motivation, emotion, affect, state, arousal, valence, positive and negative. Table 3 lists 58 studies that I found on different orders and species of mammals, in which vocalizations were analysed in relation to either arousal/valence or in relation to different contexts or situations suggesting a certain emotional arousal/valence. Variations in hunger, pain and stress were considered as similar to variations in emotional arousal. Table 3 is not exhaustive and is focused on encoding of emotions in vocalizations more than decoding. It is intended to include different orders/species and to represent biases towards certain orders/species that have been studied more than others.

Vocal correlates of arousal have been studied considerably more than correlates of valence, and most studies focused on negative situations (e.g. stress, pain, isolation, separation). Primates are the most studied order. These species often have a repertoire of several call types. Numerous studies have been conducted to investigate the contexts of production of these call variants, in order to categorize them and understand their meaning and functions (e.g. Rendall *et al.*, 1999; Scheumann *et al.*, 2007; Meise *et al.*, 2011). Some call types appear to vary gradually within and between contexts according to the caller's internal state (e.g. Coss, McCowan & Ramakrishnan, 2007). Pigs *Sus scrofa* are the most studied species, with the aim of finding vocal correlates of welfare (see also Weary & Fraser, 1995b; Weary, Ross & Fraser, 1997, not listed in Table 3).

How are vocal correlates of emotion studied?

Most studies conducted in the wild or in captivity consist in recording one or several types of vocalizations produced during naturally occurring situations characterized by different levels of arousal or variance (method = 'Observation' in Table 3). For example, Soltis, Blowers & Savage (2011) studied African elephant *Loxodonta africana* vocalizations produced during three naturally occurring social contexts; one low-arousal neutral context characterized by minimal social activity, one high-arousal negative context (dominance interaction), and one high-arousal positive context (affiliative interaction). Vocal parameters that differ between the low-arousal context (neutral) and the two high-arousal contexts (negative and positive) can be considered as indicators of arousal, whereas those that differ between the high-arousal

positive and negative situations reflect the emotional valence of the caller. Other observational studies focussed on behaviours such as dyadic agonistic interactions with low and high intensity levels in bats *Megaderma lyra* (Bastian & Schmidt, 2008), mother–pup interactions characterized by different levels of valence and arousal (reunion, separation, nursing) in Weddell seals *Leptonychotes weddellii* (Collins *et al.*, 2011) or infant restraint by female rhesus monkeys *Macaca mulatta* characterized by different threat severity levels (Jovanovic & Gouzoules, 2001). Several studies also recorded naturally occurring or experimentally elicited alarm calls, which have often been shown to simultaneously communicate the type of predator and the level of urgency (i.e. both referential and emotional information, see Manser, Seyfarth & Cheney, 2002; Seyfarth & Cheney, 2003 for a review).

Studies conducted in laboratories or on farms usually consist in placing the animals in various situations characterized by different levels of arousal or valence (method = 'Experimental' in Table 3). Most commonly, one or several types of vocalizations are recorded during complete or partial isolation or separation from conspecifics (e.g. Schrader & Todt, 1993; Byrne & Suomi, 1999; Norcross & Newman, 1999; Norcross, Newman & Cofrancesco, 1999; Yamaguchi, Izumi & Nakamura, 2010; Siebert *et al.*, 2011; Sèbe *et al.*, 2012), during human approach tests (e.g. Marchant *et al.*, 2001; Gogoleva *et al.*, 2010a,b) or during routine farm and industry-wide procedures (e.g. castration, branding; Weary *et al.*, 1998; Watts & Stookey, 1999; von Borell *et al.*, 2009). Few studies examined the relationship between vocal parameters and physiological indicators of stress (i.e. cortisol or adrenaline levels, cardiac activity; Byrne & Suomi, 1999; Norcross & Newman, 1999; Marchant *et al.*, 2001; Sèbe *et al.*, 2012). Positive vocalizations in studies investigating valence were elicited by the following situations; grooming by an experimenter (Scheumann *et al.*, 2007), friendly approach by a caretaker (Yeon *et al.*, 2011), playing (Yin & McCowan, 2004; Taylor *et al.*, 2009), feeding time (Pond *et al.*, 2010) and finally in response to a familiar companion or by activating the ascending dopaminergic system (Brudzynski, 2007).

Vocal correlates of arousal

Fifty-four of the 58 studies included in Table 3 investigated the effect of arousal on vocal parameters, making the shifts for arousal presented in Table 4 reliable. Several parameter changes were supported by more than five studies; an increase in arousal level is associated with an increase in *vocalization/element duration* and *rate*, *F0 contour*, *F0 range*, *amplitude contour*, *energy distribution* (towards higher frequencies), *peak frequency*, *formant contour*, and a decrease in *inter-vocalization/element interval*. There could also be an increase in the *number of elements* in complex vocalizations, in *H1–H2* ('hoarseness' or 'breathiness' in human voice), in *jitter*, in the *time of peak frequency* and possibly of noise (*harmonic-to-noise ratio* and *spectral noise*, but see *entropy*). Therefore, with an increase in arousal, vocalizations typically become longer, louder and harsher, with higher and more variable frequencies, and they are produced at faster rates. These changes

Table 3 Studies of vocal correlates of valence and arousal included in this review

Species		Common name	Latin name	Order	Dimension	Process	Method	Vocalization type	Parameters studied	Reference
1	2									
1	Bison	<i>Bison bison</i>	Artiodactyla	A	E	O		Antagonistic vocalizations	12	Wyman <i>et al.</i> (2008)
2	Cattle	<i>Bos taurus</i>		A	E	E		Calf vocalizations	1, 2, 5, 16	Thomas, Weary & Appleby (2001)
3				A	E	E		Distress vocalizations	1, 5, 6, 12, 16	Watts & Stookey (1999)
4	Goat	<i>Capra hircus</i>		A	E	E		Contact/distress vocalizations	1, 16, 20, 25	Siebert <i>et al.</i> (2011)
5	Pig	<i>Sus scrofa</i>		A	E	E		Distress vocalizations	1, 12, 13, 16, 17	von Borell <i>et al.</i> (2009)
6				A	E	E		Distress vocalizations	22	Düpijan <i>et al.</i> (2008)
7				A	E	E		Distress vocalizations	1, 15, 16, 20, 25	Puppe <i>et al.</i> (2005)
8				A	E	E		Distress vocalizations	1, 2, 5, 12, 15, 16, 18, 20, 21, 22, 24	Schrader & Todt (1998)
9				A	E	E		Distress vocalizations	1, 2, 16	Weary <i>et al.</i> (1998)
10				A	E/D	E		Distress vocalizations	1, 2, 15, 16	Weary & Fraser (1995a)
11				A	E	E		Short/long grunts and squeals	1, 2	Marchant <i>et al.</i> (2001)
12	Sheep	<i>Ovis aries</i>		A	E	E		Contact/distress vocalizations	1, 5, 12, 15, 20	Sébe <i>et al.</i> (2012)
13	Cat	<i>Felis catus</i>	Carnivora	AV	E	E		Antagonistic and contact vocalizations	1, 2, 5, 15, 16, 22	Yeon <i>et al.</i> (2011)
14	Dog	<i>Canis familiaris</i>		V	E/D	E		Growls and barks	1, 3, 5, 21, 23	Taylor <i>et al.</i> (2009)
15				AV	E	E		Barks	1, 3, 12, 13, 16, 17, 18, 20, 21, 26	Yin & McCowan (2004)
16	Mongoose	<i>Suricata suricatta</i>		A	E	E/O		Alarm vocalizations	1, 2, 3, 15, 16, 17, 18, 20, 21, 24	Manser (2001)
17	Silver fox	<i>Vulpes vulpes</i>		A	E	E		Antagonistic vocalizations	1, 2, 15, 16, 25	Gogoleva <i>et al.</i> (2010b)
18				AV	E	E		Antagonistic and contact vocalizations	1, 2, 16	Gogoleva <i>et al.</i> (2010a)
19	Spotted hyena	<i>Crocuta crocuta</i>		A	E	O		Whoops	1, 2, 3, 5	Theis <i>et al.</i> (2007)
20	Weddell seal	<i>Leptonychotes weddellii</i>		AV	E	O		Contact vocalizations	1, 2, 5, 15, 21	Collins <i>et al.</i> (2011)
21	Bottlenose dolphin	<i>Tursiops truncatus</i>	Cetacea	A	E	E		Contact vocalizations	1, 2, 3, 4, 5	Esch <i>et al.</i> (2009)
22	Greater false vampire bat	<i>Megaderma lyra</i>	Chiroptera	A	E	O		Antagonistic and response vocalizations	1, 2, 3, 4, 5, 8, 16	Bastian & Schmidt (2008)
23	Horse	<i>Equus caballus</i>	Perissodactyla	V	E	E		Whinnies	Hidden Markov Model	Pond <i>et al.</i> (2010)

Table 3 Continued.

Species		Order		Dimension	Process	Method	Vocalization type	Parameters studied	Reference
Common name	Latin name	Order							
24	Barbary macaque	<i>Macaca sylvanus</i>	Primates	A	E/D	E/O	Disturbance call	1, 3, 12, 16, 17, 18, 20, 21	Fischer, Hammerschmidt, & Todt (1995)
25	Bonnet macaque	<i>Macaca radiata</i>		A	E	E	Alarm vocalizations	1, 5, 6, 7, 14, 21, 24, 26	Coss <i>et al.</i> (2007)
26	Baboon	<i>Papio cynocephalus ursinus</i>		A	E	O	Grunts	1, 3, 5, 7, 8, 9, 15, 16, 17, 18, 20, 22, 24, 26	Meise <i>et al.</i> (2011)
27				A	E	O	Contact and alarm vocalizations	1, 2, 5, 15, 16, 17, 18, 19, 21, 22, 23, 24	Fischer <i>et al.</i> (2001)
28				A	E	O	Grunts	1, 2, 3, 5, 8, 10, 12, 19, 22	Rendall (2003)
29				A	E	O	Grunts	1, 5, 15, 19, 22	Owren <i>et al.</i> (1997)
30	Chimpanzee	<i>Pan troglodytes</i>		A	E	O	Antagonistic vocalizations	1, 3, 18, 21	Siebert & Parr (2003)
31				A	E	O	Antagonistic vocalizations	1, 2, 5, 16, 17, 22	Slocombe & Zuberbühler (2007)
32	Common marmoset	<i>Callithrix jacchus</i>		A	E	E	Contact vocalizations	1, 4, 5, 6, 12, 21	Schrader & Todt (1993)
33				A	E	E	Contact/distress vocalizations	2	Norcross & Newman (1999)
34				A	E	E	Contact/distress vocalizations	1, 3, 5, 6, 7	Norcross <i>et al.</i> (1999)
35				A	E	E	Contact/distress vocalizations	1, 2, 3, 5, 7, 12, 16, 17	Yamaguchi <i>et al.</i> (2010)
36	Gray mouse lemur	<i>Microcebus murinus</i>		AV	E	E	Whistle, tsak and purr	1, 3, 5, 16, 20	Scheumann <i>et al.</i> (2007)
37	Japanese macaque	<i>Macaca fuscata</i>		A	E	O	Coo calls	1, 3, 5, 6, 7, 17	Sugiura (2007)
38	Pigtail macaque	<i>Macaca nemestrina</i>		A	E	O	Antagonistic vocalizations	1, 5, 16, 17, 18, 20, 21, 24	Gouzoules & Gouzoules (1989)
39	Redfronted lemur	<i>Eulemur fulvus rufus</i>		A	E/D	E/O	Alarm vocalizations	12, 15, 16, 18, 20, 22	Fichtel & Hammerschmidt (2002)
40	Squirrel monkey	<i>Saimiri sciureus</i>		A	D	E	Alarm vocalizations	12, 15, 16	Fichtel & Hammerschmidt (2003)
41				AV	E	E	Eight different aversive/rewarding call types	1, 5, 15, 16, 18, 20, 24	Fichtel <i>et al.</i> (2001)
42	Thomas's langur	<i>Presbytis thomasi</i>		A	E/D	O	Loud calls	1, 3, 4, 5	Wich <i>et al.</i> (2009)
43	Tufted capuchin	<i>Cebus apella</i>		A	E	E	Distress vocalizations	2	Byrne & Suomi (1999)
44	Rhesus monkey	<i>Macaca mulatta</i>		AV	E	O	Infant vocalizations	2	Jovanovic & Gouzoules (2001)

Table 3 Continued.

Species		Latin name	Order	Dimension	Process	Method	Vocalization type	Parameters studied	Reference
Common name									
45	Rhesus monkey/African elephant	<i>Macaca mulatta/Loxodonta africana</i>	Primates/ Proboscidea	A	E	O	Rumbles/infant vocalizations	10, 11	Li <i>et al.</i> (2007)
46	African elephant	<i>Loxodonta africana</i>	Proboscidea	A	E	O	Rumbles	1, 5, 6, 7, 10, 15, 22	Soltis, Leong & Savage (2005)
47				A	E	O	Distress vocalizations	1, 5, 16, 24, 26, 27	Stoeger <i>et al.</i> (2011)
48				AV	E	O	Rumbles	1, 5, 6, 8, 12, 13, 22	Soltis <i>et al.</i> (2011)
49				A	E	O	Rumbles	1, 5, 6, 10, 11, 12, 22, 23, 26	Soltis <i>et al.</i> (2009)
50	Alpine marmot	<i>Marmota marmota</i>	Rodentia	A	E/D	E/O	Alarm vocalizations	1, 2, 3, 4, 5, 16, 20	Blumstein & Arnold (1995)
51	Guinea pig	<i>Cavia porcellus</i>		A	E	E	Distress vocalizations	1, 3, 5, 16, 21	Monticelli <i>et al.</i> (2004)
52	Rat	<i>Rattus norvegicus</i>		V	E	E	Ultrasonic vocalizations	1, 16, 20	Brudzynski (2007)
53				V	D	E	Ultrasonic vocalizations	16	Burman <i>et al.</i> (2007)
54	Yellow-bellied marmot	<i>Marmota flaviventris</i>		A	D	E	Alarm vocalizations	24	Blumstein & Récapet (2009)
55				A	E/D	E/O	Alarm vocalizations	1, 2, 3, 5, 6, 12, 16	Blumstein & Armitage (1997)
56				A	E	E/O	Alarm vocalizations	1, 25	Blumstein & Chi (2011)
57	Tree shrew	<i>Tupaia belangeri</i>	Scandentia	A	E	E	Antagonistic vocalizations	1, 3, 4, 5, 16	Schehka, Esser & Zimmermann (2007)
58				A	E	E	Disturbance calls	1, 3, 4, 5, 12, 16, 22	Schehka & Zimmermann (2009)

Vocalization type: category of vocalization studied (when the vocalization(s) studied could be emitted in various contexts, the original name is given instead of the category). Parameter studied: vocal parameters or categories of parameters measured (the numbers correspond to the codes given in Table 1). Only the parameters listed in Table 1 (i.e. commonly used across studies) are listed. Parameter names differed across studies. The names listed in Table 1 might not correspond to the exact name used in the reference cited, but correspond to the description given by the authors.

Dimension: A, arousal; V, valence. Process: E, encoding; D, decoding. Method: E, experimental; O, observation.

Table 4 Changes in vocal parameters according to arousal and valence

Category	Parameter	Arousal (low to high)		Valence (negative to positive)	
			Evidence		Evidence
Time parameters	Vocalization/element duration	◇	<:5,7,9,10,11,12,15,20,22,24,26,28,31,32,35,37,38,41,47,48,49 >:8,16,27,30,34,51,56,58	>	>:14,41,52
	Vocalization/element rate	◇	<:2,9,10,11,13,16,17,18,19,20,21,22,33,35,43,44,55 >:31	–	–
	Inter-vocalization/element interval	◇	>:15,16,19,22,24,26,28,30,35,55,57,58 <:34	>	>:14
	Number of elements	◇	<:21,22,42,58 >:32,50	–	–
<i>F0</i>	<i>F0</i> contour	<	<:2,3,8,12,13,19,20,21,22,28,31,32,34,37,41,48,49,51,55,57,58	◇	>:41,48 <:14
	<i>F0</i> range	<	<:3,32,34,37,48,55	>	>:48
	Time of maximum/minimum <i>F0</i>	<	<:35	–	–
	H1–H2	<	<:28,48	>	>:48
	<i>F0</i> slope	–	–	–	–
	Jitter	<	<:45,46,49	–	–
	Shimmer	<	<:45	–	–
	Amplitude contour	<	<:1,3,5,8,12,32,35,39,40,48,49,55	–	–
Amplitude	Amplitude range	◇	<:48 >:15	<	<:15
	AM	–	–	–	–
Frequency spectrum	Energy distribution	◇	<:13,17,20,39,40,41 >:12,26	◇	<:13 >:41
	Peak frequency	<	<:7,8,13,17,27,31,35,38,40,41,58	◇	<:13,52,53 >:18,41
	Time of peak frequency	<	<:26,27,31,35	–	–
	Dominant frequency band contour	◇	<:24,27,41 >:15,26	>	>:41
	Spectral slope	–	–	–	–
	Frequency range	◇	<:8,26,38,41 >:4,7,12,15	◇	<:52 >:41
	FM	◇	<:8,24,38 >:27,30,32	>	>:14
	<i>F_n</i> contour (<i>F1</i> , <i>F2</i> , <i>F3</i> , etc.)	◇	<:6,27,29,31,48,58 >:26,28	<	<:13
Non-linear phenomena	Formant dispersion	–	–	–	–
	Spectral noise	◇	<:16,27,41,54 >:38	>	>:41
	Entropy	◇	>:4,7,56 <:17	–	–
	Harmonic-to-noise ratio	>	>:15,25,26,47	–	–
	Subharmonics	–	–	–	–

Arousal: '<' indicates an increase in parameter value with an increase in arousal; '>' indicates a decrease in parameter value with an increase in arousal. Valence: '<' indicates that parameter value is higher in the positive than negative situation; '>' indicates that parameter value is lower in the positive than negative situation; '–' indicates that no study has found a significant shift for this parameter. For *F0* slope, '<' indicates a steeper slope. For energy distribution, '<' indicates a shift in energy distribution towards higher frequencies. The Evidence column lists which studies found an increase or decrease in parameters. The numbers for the references corresponds to the codes given in Table 3. This table only includes studies that reported either a) a significant within-call type difference in parameters between situations or b) a significant difference in the proportion of call types produced in various situations associated with a significant difference in parameters between the call types considered. AM, amplitude modulation; FM, frequency modulation.

correspond closely to those described for humans (Scherer, 1986; Murray & Arnott, 1993; Bachorowski & Owren, 1995; Banse & Scherer, 1996; Zei Pollermann & Archinard, 2002; Juslin & Scherer, 2005; Li *et al.*, 2007). Furthermore, they correspond closely to the effects of the physiological changes linked to an increase in arousal on the acoustic structure of vocalizations, which have been described in humans (Scherer, 1986); increase in the action and/or tension of the respiratory muscles (longer *duration*, higher *amplitude* and higher *F0*), decrease in salivation (higher *formant frequencies*), increase in the action and/or tension of the cricothyroid muscles that stretch the vocal folds (higher *F0*), and increase in pharyngeal constriction and tension of the vocal tract walls (increase of the proportion of energy in the upper part of the *frequency spectrum*). The other parameter changes listed in Table 4 are supported by only one study or are not clear (i.e. both increases and decreases have been reported).

There is strong evidence for the increase in arousal level associated with the increase in *vocalization/element rate*, *F0 contour*, *F0 range*, *amplitude contour*, *energy distribution* (towards higher frequencies), *frequency peak* and *formant contour* and the decrease in *inter-vocalization interval* (5–21 studies, maximum two studies with opposite shift). These parameters appear therefore as ideal indicators of arousal. By contrast, the increase in *vocalization/element duration* is challenged by eight studies. For example, the increase in *duration* was not found for some alarm calls (Manser, 2001; Blumstein & Chi, 2011). In meerkats *Suricata suricatta*, for a given class of predator, high-urgency situations seem to elicit longer calls than low-urgency situations. However, shorter alarm calls are given in response to more dangerous predators compared with distant predators or non-dangerous animals (Manser, 2001). Similarly, Blumstein & Arnold (1995) found that Alpine marmots *Marmota marmo* produce alarm calls with fewer elements in higher-urgency situations. Shorter alarm calls may reduce conspicuousness to predators and allow a faster response. *Duration* also decreased in guinea pigs *Cavia porcellus* with presumed higher arousal levels during periods of isolation (Monticelli, Tokumaru & Ades, 2004). In the same way, in piglets, the initial increase in *duration* and in most of the vocal parameters during the first 2 min of isolation was followed by a decrease (Weary & Fraser, 1995a). These changes are most likely linked to a decrease in motivation, independently of stress (Monticelli *et al.*, 2004). Therefore, motivation levels should be taken into account in the interpretation of context-related changes in vocal parameters.

In the case of *non-linear phenomena*, the results are not consistent. According to Table 4, *harmonic-to-noise ratio* decreases, *spectral noise* increases (more noise), but entropy decreases (more pure tone vocalizations) with arousal. The increase in *spectral noise* (Table 4) is contradicted by Gouzoules & Gouzoules (1989), which showed that pigtail macaques *Macaca nemestrina* produced less noisy and more tonal screams during contact aggression (high presumed arousal) than during non-contact aggression. Similarly, Blumstein & Chi (2011) showed that yellow-bellied marmots *Marmota flaviventris* with more faecal glucocorticoid metabolites, indicating higher stress levels, produced less noisy calls

(measured as *entropy*). Therefore, it seems that non-linear phenomena might increase or decrease with arousal depending on species or particular contexts and are not good indicators of arousal.

Vocal correlates of valence

Vocal correlates of valence have been considerably less studied than arousal (Table 4). There are only a few studies in which authors compared vocalizations produced in negative and positive situations. There are two main reasons for this lack of research. One is the difficulty to find calls that are produced in positive situations (but see exceptions of positive vocalizations later). Because vocal correlates of negative states signal urgency (e.g. alarm calls) and need (e.g. infant begging calls), these vocalizations are far more common than positive vocalizations, and probably emerged earlier during evolution. The evolution of positive vocalizations could have been facilitated later by the increased importance of communication within social groups (Brudzynski, 2007). Expression of arousal can be studied by comparing vocalizations produced in negative situations that are characterized by varying degrees of arousal. By contrast, research on expression of valence must compare vocalizations produced in positive and negative situations that are characterized by a similar degree of arousal. This leads to the second reason for a lack of research on vocal correlates of valence; it is difficult to find situations of opposite valence, but similar arousal. Expressions of negative emotions (e.g. physiological, visual, vocal) are easier to study, because they are often more intense than expressions of positive emotions (Boissy *et al.*, 2007). Therefore, it is difficult to find situations triggering positive emotions as intense as negative emotions.

Because of this lack of research, knowledge on vocal correlates of valence listed in Table 4 is sparse. Some studies show a shift towards higher frequencies during positive situations. In dogs, barks emitted during positive situations (play) are characterized by wider *amplitude ranges*, shorter *inter-call intervals*, shorter *durations*, higher *F0* and smaller *frequency modulations* compared with barks emitted in negative situations of probably similar arousal (Yin & McCowan, 2004; Taylor *et al.*, 2009). Yeon *et al.* (2011) showed that feral cats *Felis catus* produce vocalizations with higher *energy distributions*, *F1* and *peak frequencies* in affiliative compared with agonistic situations. However, the 'affiliative' situation in this case was an approach by a familiar caretaker, and it is not clear how positive or intense this experience was for feral cats. Pond *et al.* (2010) found spectral differences between vocalizations produced in two situations of similar arousal and different valence using Hidden Markov Models, but the shifts in individual vocal parameters are not detailed in this study.

There is also evidence for a shift towards low frequencies during positive situations. Jovanovic & Gouzoules (2001) and Scheumann *et al.* (2007) showed that infant Rhesus monkeys and gray mouse lemurs produce different kinds of calls during positive contexts ('coos' and 'purr' respectively) compared with negative contexts. 'Coos' and 'purr' are both characterized by low frequencies. Fichtel, Hammerschmidt & Jürgens (2001) found that in squirrel monkeys *Saimiri sciureus*, call

level of 'negativity' (aversion) is generally correlated with longer *duration*, higher *F0* *contour*, *energy distribution*, *peak frequency*, *dominant frequency band contour*, wider *frequency range*, and more *noise*. However, it is not clear how much of this variance is explained by arousal or valence. Tame and aggressive silver foxes *Vulpes vulpes* differ in their reactions to humans; tame foxes show a decrease and aggressive foxes an increase in *peak frequency* during approach (Gogoleva *et al.*, 2010a), suggesting that low-*peak frequencies* reflects positive emotions. Soltis *et al.* (2011) found that African elephant rumbles produced in a positive situation have lower *F0*, *H1–H2* and narrower *F0 range* than those produced in a negative situation. However, because the shifts in these parameters occurring between the neutral and positive contexts were similar (i.e. same direction), yet less intense, than the shifts exhibited between the neutral and negative contexts, the authors suggested that their results were more consistent with an effect of emotional arousal than valence. Similarly, the variations between contexts in vocal parameters found by Collins *et al.* (2011) in Weddell seals were more consistent with the expression of emotional arousal. Therefore, the only parameter shift that is supported by three studies, without any opposite shift, is *duration*, with positive situations characterized by shorter vocalizations (Table 4).

There are some good examples in the literature of vocal expression of positive emotions: purr, laughter and rat ultrasonic 50-Hz vocalizations. Felid purrs are low pitched vocalizations (mean *F0* = 26.3 Hz), characterized by a pulse-train structure and low *amplitude*, and produced more or less continuously for up to several minutes (Peters, 2002). They can be mixed with other tonal vocalizations (e.g. meow in cats) produced at the same time (McComb *et al.*, 2009). Vocalizations that are structurally similar to purring have also been reported in several Carnivora families and other mammals, including primates (e.g. ring-tailed lemur *Lemur catta*, Macedonia, 1993; tree shrew *Tupaia belangeri*; Benson, Binz & Zimmermann, 1976). Purring is produced mostly by juveniles, but also by adults, in positive contexts (relaxed, friendly) such as nursing/suckling, mutual grooming, courtship or friendly approach (Peters, 2002). The wide distribution of purring-like vocalizations among mammals shows that vocalizations produced in 'friendly' contexts do not always comply with the predicted motivation-structural rules (i.e. expecting high, pure tone-like sounds in friendly contexts; Morton, 1977).

Human laughter is another well-known positive vocalization. Laughter consists of a repetition of vowel-like bursts (fricative, i.e. aspired 'h' sound, followed by a vowel). It is characterized by a high *F0*, on average twice higher than in modal speech (282 Hz vs. 120 Hz for men, and 421 Hz vs. 220 Hz for women; Bachorowski, Smoski & Owren, 2001). Other characteristics of laughter include a salient *F0 modulation*, high *F1* compared with normal speech vowels because of wide jaw opening and pharyngeal constriction, and the presence of *non-linear phenomena* (e.g. subharmonics and biphonation; Bachorowski *et al.*, 2001; Szameitat *et al.*, 2011). Young orangutans *Pongo pygmaeus*, gorillas *Gorilla gorilla*, chimpanzees, bonobos *P. paniscus* and siamang *Symphalangus syndactylus* produce very similar vocalizations, mostly

noisy, that can be elicited by tickling, suggesting that 'laughter' is a cross-species phenomenon (Ross, Owren & Zimmermann, 2009).

Rats produce two types of ultrasonic vocalizations, 22- and 50-kHz vocalizations. There is substantial evidence from ethological, pharmacological, and brain stimulation studies that these two types of calls reflect the emotional valence of the caller, either negative (22 kHz alarm calls) or positive (50 kHz social calls, e.g. Knutson *et al.*, 2002; Burgdorf & Moskal, 2009). Vocalizations of 22 kHz are typically produced during anticipation of punishment or avoidance behaviour, whereas 50 kHz vocalizations occur during anticipation of reward or approach behaviour. Vocalizations of 50 kHz are emitted particularly during play, and can also be produced in response to manual tickling by an experimenter (Panksepp & Burgdorf, 2000). Therefore, they have been suggested to be a primal form of laughter (Panksepp & Burgdorf, 2003; Panksepp, 2009). Rat ultrasonic vocalizations have been linked to neural substrates responsible for negative and positive states (ascending cholinergic and dopaminergic systems; Brudzynski, 2007). Negative vocalizations are characterized by longer *durations*, lower-*peak frequencies* and narrower *frequency ranges* (bandwidth) than positive ones (Brudzynski, 2007). These structural differences between these two call types are perceived by receivers and induce different behaviours suggesting negative (22 kHz) or positive (50 kHz) internal states (Burman *et al.*, 2007).

To summarize, vocalizations produced in positive situations could be shorter in *duration*, but seem to vary in *F0*, from very low 'purr' in felids to high-frequency 50-kHz vocalizations in rats and laughter in humans. More parameters need to be investigated to find vocal correlates of valence in animals. For example, in humans, positive emotions are characterized by a lower *amplitude*, shifts in the *energy distribution* towards low frequencies, an earlier position of the *maximum peak frequency*, narrower *frequency ranges*, steeper *spectral slope*, higher *formants* and less *spectral noise* (Zei Pollermann & Archinard, 2002; Waaramaa *et al.*, 2006, 2010; Hammer-schmidt & Jürgens, 2007; Goudbeek & Scherer, 2010). These parameters might also express valence in other mammals.

Summary of evidence

Vocal expression of arousal has been extensively studied. The best indicators of arousal have been *vocalization/element rate*, *F0 contour*, *F0 range*, *amplitude contour*, *energy distribution*, *frequency peak* and *formant contour* (increase with arousal) and *inter-vocalization interval* (decreases with arousal). Because of a lack of research on the topic, no clear indicator of valence has been found yet. Likely candidates include indicators of valence found in humans, such as *amplitude level*, *energy distribution*, *maximum peak frequency*, *frequency range*, *spectral slope*, *formants* and *spectral noise*. In particular, *formant* parameters are rarely measured in humans and in other animals (Scherer, 2003; Juslin & Scherer, 2005). Several studies suggested that this could be the key to the vocal differentiation of emotional valence (Scherer, 1986; Banse & Scherer, 1996; Waaramaa *et al.*, 2010; Patel *et al.*, 2011).

Humans benefit from enhanced motor control and flexibility of the vocal articulators (tongue, lips, velum, jaw, etc.), allowing us to create different patterns of changes in *F1* and *F2* (Fant, 1960). Other species of mammals have a smaller degree of flexibility in vocal tract length and shape, and therefore less possibility to alter *formant frequencies*. However, variation in vocal tract length can be achieved by various mechanisms including lips extension, modification of the level of nasalization, and most commonly, retraction of the larynx into the throat (Owren, Seyfarth & Cheney, 1997; Fitch, 2000b; Fitch & Reby, 2001; Harris *et al.*, 2006; McElligott, Birrer & Vannoni, 2006). Indicators of emotional valence would be particularly useful for assessing animal welfare (Manteuffel *et al.*, 2004). For example, vocal cues to positive emotions could enhance positive welfare, i.e. promote positive experiences in captive animals (Boissy *et al.*, 2007).

Research on vocal correlates of emotions needs to make clearer assumptions regarding the emotion triggered by the observed or experimental situation. The emotional valence and arousal elicited by the situation could be verified using other components of emotions, like physiological indicators (e.g. cortisol or adrenaline levels, cardiac activity; Byrne & Suomi, 1999; Norcross & Newman, 1999; Marchant *et al.*, 2001; Sèbe *et al.*, 2012). In natural settings, several behavioural indicators of emotions can be used (see Schehka & Zimmermann, 2009; Zimmermann, 2009; Stoeger *et al.*, 2011). Studies on vocal correlates of arousal should focus on vocalizations recorded during situations characterized by different levels of arousal and a similar valence, whereas studies on vocal expression of valence should investigate vocalizations recorded during situations characterized by opposite valences (positive and negative) and a similar arousal level. When possible, studies should focus on one given type of vocalization and measure its variation between contexts, instead of investigating differences between call types produced in various contexts. Finally, calls vary according to states other than emotions, such as motivation (e.g. aversion, attraction; Morton, 1977; August & Anderson, 1987; Ehret, 2006), which could be taken into account when interpreting context-related vocal variation, in the same way as the potency dimension (i.e. level of control of the situation) used in studies on affective prosody (Juslin & Scherer, 2005).

Conclusions

This review shows that the increase in *vocalization/element rate*, *F0 contour*, *F0 range*, *amplitude contour*, *energy distribution*, *frequency peak* and *formant contour* and the decrease in *inter-vocalization interval* are particularly good indicators of arousal. By contrast, indicators of valence still need to be investigated. In humans, as in other mammals, expression and perception of emotion is crucial to regulate social interactions. A deficit in either expression or perception can result in profound deficits in social relationships (Bachorowski, 1999). The general interest in the field of animal emotion is growing quickly, and is relevant to several disciplines such as evolutionary zoology, affective neuroscience, comparative psychology, animal welfare science and psychopharmacology (Mendl

et al., 2010). Because the subjective component of emotional experiences are not yet possible to prove or measure in animals, other indicators are needed to infer emotional states (e.g. neurophysiological, behavioural and/or cognitive). In particular, indicators of positive emotions are lacking (Boissy *et al.*, 2007). Vocal indicators of emotions in animals could represent convenient and non-invasive indicators, which would be particularly useful to assess and improve welfare (Weary & Fraser, 1995b; Watts & Stookey, 2000; Manteuffel *et al.*, 2004; Schön, Puppe & Manteuffel, 2004). Findings on vocal correlates of emotions in mammals could also serve as a useful model for studies on humans, in which a greater motor control results in confounding factors influencing affective prosody (Scherer *et al.*, 1984; Scheiner & Fisher, 2011).

Acknowledgements

I am grateful to Alan McElligott, Megan Wyman, Anna Taylor and an anonymous referee for helpful comments on the manuscript. I acknowledge the financial support of the Swiss National Science Foundation and the Swiss Federal Veterinary Office.

References

- Adolphs, R. (2010). Emotion. *Curr. Biol.* **20**, R549–552.
- August, P.V. & Anderson, J.G.T. (1987). Mammal sounds and motivation – structural rules: a test of the hypothesis. *J. Mammal.* **68**, 1–9.
- Bachorowski, J.-A. (1999). Vocal expression and perception of emotion. *Curr. Dir. Psychol. Sci.* **8**, 53–57.
- Bachorowski, J.-A. & Owren, M.J. (1995). Vocal expression of emotion: acoustic properties of speech are associated with emotional intensity and context. *Psychol. Sci.* **6**, 219–224.
- Bachorowski, J.-A., Smoski, M.J. & Owren, M.J. (2001). The acoustic features of human laughter. *J. Acoust. Soc. Am.* **110**, 1581–1597.
- Banse, R. & Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* **70**, 614–636.
- Bastian, A. & Schmidt, S. (2008). Affect cues in vocalizations of the bat, *Megaderma lyra*, during agonistic interactions. *J. Acoust. Soc. Am.* **124**, 598–608.
- Benson, B.N., Binz, H. & Zimmermann, E. (1976). Vocalizations of infant and developing tree shrews (*Tupaia belangeri*). *J. Mammal.* **41**, 80–106.
- van Bezooijen, R. (1984). *The characteristics and recognizability of vocal expression of emotions*. Dordrecht: Foris.
- Blumstein, D.T. & Armitage, K.B. (1997). Alarm calling in yellow-bellied marmots: I. The meaning of situationally variable alarm calls. *Anim. Behav.* **53**, 143–171.
- Blumstein, D.T. & Arnold, W. (1995). Situational specificity in Alpine-marmot alarm communication. *Ethology* **100**, 1–13.
- Blumstein, D.T. & Chi, Y.Y. (2011). Scared and less noisy: glucocorticoids are associated with alarm call entropy. *Biol. Lett.* **8**, 189–192.

- Blumstein, D.T. & Récapet, C. (2009). The sound of arousal: the addition of novel non-linearities increases responsiveness in marmot alarm calls. *Ethology* **115**, 1074–1081.
- Boissy, A., Manteuffel, G., Jensen, M.B., Moe, R.O., Spruijt, B., Keeling, L.J., Winckler, C., Forkman, B., Dimitrov, I., Langbein, J., Bakken, M., Veissier, I. & Aubert, A. (2007). Assessment of positive emotions in animals to improve their welfare. *Physiol. Behav.* **92**, 375–397.
- von Borell, E., Büniger, B., Schmidt, T. & Horn, T. (2009). Vocal-type classification as a tool to identify stress in piglets under on-farm conditions. *Anim. Welf.* **18**, 407–416.
- von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant-Forde, J., Marchant-Forde, R., Minero M., Mohr E., Prunier A., Valance D. & Veissier I. (2007). Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals – a review. *Physiol. Behav.* **92**, 293–316.
- Briefer, E. & McElligott, A.G. (2011). Indicators of age, body size and sex in goat kid calls revealed using the source-filter theory. *Appl. Anim. Behav. Sci.* **133**, 175–185.
- Briefer, E., Vannoni, E. & McElligott, A.G. (2010). Quality prevails over identity in the sexually selected vocalisations of an ageing mammal. *BMC Biol.* **8**, 35.
- Brudzynski, S.M. (2007). Ultrasonic calls of rats as indicator variables of negative or positive states: acetylcholine–dopamine interaction and acoustic coding. *Behav. Brain Res.* **182**, 261–273.
- Burgdorf, J. & Moskal, J.R. (2009). Frequency modulated 50 kHz ultrasonic vocalizations reflect a positive emotional state in the rat: neural substrates and therapeutic implications. In *Handbook of mammalian vocalization – an integrative neuroscience approach*: 209–214. Brudzynski, S.M. (Ed.). London: Academic Press.
- Burgdorf, J., Panksepp, J. & Moskal, J.R. (2011). Frequency-modulated 50 kHz ultrasonic vocalizations: a tool for uncovering the molecular substrates of positive affect. *Neurosci. Biobehav. Rev.* **35**, 1831–1836.
- Burgdorf, J., Wood, P.L., Kroes, R.A., Moskal, J.R. & Panksepp, J. (2007). Neurobiology of 50-kHz ultrasonic vocalizations in rats: electrode mapping, lesion, and pharmacology studies. *Behav. Brain Res.* **182**, 274–283.
- Burman, O.H.P., Ilyat, A., Jones, G. & Mendl, M. (2007). Ultrasonic vocalizations as indicators of welfare for laboratory rats (*Rattus norvegicus*). *Appl. Anim. Behav. Sci.* **104**, 116–129.
- Byrne, G. & Suomi, S.J. (1999). Social separation in infant *Cebus apella*: patterns of behavioral and cortisol response. *Int. J. Dev. Neurosci.* **17**, 265–274.
- Charlton, B.D., Keating, J.L., Kersey, D., Rengui, L., Huang, Y. & Swaisgood, R.R. (2011). Vocal cues to male androgen levels in giant pandas. *Biol. Lett.* **7**, 71–74.
- Collias, N.E. (1960). An ecological and functional classification of animals sounds. In *Animal sounds and communication*: 368–391. Lanyon, W.E. & Tavolga, W.N. (Eds). Washington, DC: American Institute of Biological Science.
- Collins, K.T., McGreevy, P.D., Wheatley, K.E. & Harcourt, R.G. (2011). The influence of behavioural context on Weddell seal (*Leptonychotes weddellii*) airborne mother-pup vocalisation. *Behav. Proc.* **87**, 286–290.
- Compton, L.A., Clarke, J.A., Seidensticker, J. & Ingrisano, D.R. (2001). Acoustic characteristics of white-nosed coati vocalizations: a test of motivation-structural rules. *J. Mammal.* **82**, 1054–1058.
- Coss, R.G., McCowan, B. & Ramakrishnan, U. (2007). Threat-related acoustical differences in alarm calls by wild bonnet macaques (*Macaca radiata*) elicited by python and leopard models. *Ethology* **113**, 352–367.
- Darwin, C. (1872). *The expressions of emotions in man and animals*. London: John Murray.
- Davidson, R.J., Scherer, K.R. & Goldsmith, H.H. (2003). *Handbook of affective sciences*. Oxford: Oxford University Press.
- Dawkins, M.S. (2008). The science of animal suffering. *Ethology* **114**, 937–945.
- de Waal, F.B.M. (2011). What is an animal emotion? *Ann. N. Y. Acad. Sci.* **1224**, 191–206.
- Désiré, L., Boissy, A. & Veissier, I. (2002). Emotions in farm animals: a new approach to animal welfare in applied ethology. *Behav. Proc.* **60**, 165–180.
- Düppjan, S., Schön, P.-C., Puppe, B., Tuchscherer, A. & Manteuffel, G. (2008). Differential vocal responses to physical and mental stressors in domestic pigs (*Sus scrofa*). *Appl. Anim. Behav. Sci.* **114**, 105–115.
- Ehret, G. (2006). Common rules of communication soundperception. In *Behaviour and neurodynamics for auditory communication*: 85–114. Kanwal, J. & Ehret, G. (Eds). Cambridge: Cambridge University Press.
- Esch, H.C., Sayigh, L.S., Blum, J.E. & Wells, R.S. (2009). Whistles as potential indicators of stress in bottlenose dolphins (*Tursiops truncatus*). *J. Mammal.* **90**, 638–650.
- Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.
- Feighny, J.A., Williamson, K.E. & Clarke, J.A. (2006). North American elk bugle vocalizations: male and female bugle call structure and context. *J. Mammal.* **87**, 1072–1077.
- Fichtel, C. & Hammerschmidt, K. (2002). Responses of red-fronted lemurs to experimentally modified alarm calls: evidence for urgency-based changes in call structure. *Ethology* **108**, 763–777.
- Fichtel, C. & Hammerschmidt, K. (2003). Responses of squirrel monkeys to their experimentally modified mobbing calls. *J. Acoust. Soc. Am.* **113**, 2927–2932.
- Fichtel, C., Hammerschmidt, K. & Jürgens, U. (2001). On the vocal expression of emotion. A multi-parametric analysis of different states of aversion in the squirrel monkey. *Behaviour* **138**, 97–116.
- Fischer, J., Hammerschmidt, K., Cheney, D.L. & Seyfarth, R.M. (2001). Acoustic features of female chacma baboon barks. *Ethology* **107**, 33–54.

- Fischer, J., Hammerschmidt, K. & Todt, D. (1995). Factors affecting acoustic variation in Barbary-macaque (*Macaca sylvanus*) disturbance calls. *Ethology* **101**, 51–66.
- Fitch, W.T. (2000a). The evolution of speech: a comparative review. *Trends Cogn. 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. & Hauser, M.D. (1995). Vocal production in non-human primates: acoustics, physiology, and functional constraints on 'honest' advertisement. *Am. J. Primatol.* **37**, 191–219.
- Fitch, W.T. & Reby, D. (2001). The descended larynx is not uniquely human. *Proc. R. Soc. Lond. Ser. B* **268**, 1669–1675.
- Frolov, M.V., Milovanova, G.B., Lazarev, N.V. & Mekhedova, A.Y. (1999). Speech as an indicator of the mental status of operators and depressed patients. *Hum. Physiol.* **25**, 42–47.
- Gogoleva, S., Svetlana, S., Volodin, I., Volodina, E., Kharlamova, A. & Trut, L.N. (2010a). Sign and strength of emotional arousal: vocal correlates of positive and negative attitudes to humans in silver foxes (*Vulpes vulpes*). *Behaviour* **147**, 1713–1736.
- Gogoleva, S., Volodina, E., Volodin, I., Kharlamova, A. & Trut, L. (2010b). The gradual vocal responses to human-provoked discomfort in farmed silver foxes. *Acta Ethol.* **13**, 75–85.
- Goudbeek, M. & Scherer, K.R. (2010). Beyond arousal: valence and potency/control cues in the vocal expression of emotion. *J. Acoust. Soc. Am.* **128**, 1322–1336.
- Gouzoules, H. & Gouzoules, S. (1989). Design features and developmental modification of pigtail macaque, *Macaca nemestrina*, agonistic screams. *Anim. Behav.* **37**, 383–401.
- Gouzoules, H. & Gouzoules, S. (2000). Agonistic screams differ among four species of macaques: the significance of motivation-structural rules. *Anim. Behav.* **59**, 501–512.
- Grandin, T. (2001). Cattle vocalizations are associated with handling and equipment problems at beef slaughter plants. *Appl. Anim. Behav. Sci.* **71**, 191–201.
- Gros-Louis, J., Perry, S., Fichtel, C., Wikberg, E., Gilkenson, H., Wofsy, S. & Fuentes, A. (2008). Vocal repertoire of *Cebus capucinus*: acoustic structure, context, and usage. *Int. J. Primatol.* **29**, 641–670.
- Hammerschmidt, K. & Jürgens, U. (2007). Acoustical correlates of affective prosody. *J. Voice* **21**, 531–540.
- Harris, T.R., Fitch, W.T., Goldstein, L.M. & Fashing, P.J. (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.
- Jovanovic, T. & Gouzoules, H. (2001). Effects of nonmaternal restraint on the vocalizations of infant rhesus monkeys (*Macaca mulatta*). *Am. J. Primatol.* **53**, 33–45.
- Jürgens, U. (2009). The neural control of vocalization in mammals: a review. *J. Voice* **23**, 1–10.
- Juslin, P. & Scherer, K.R. (2005). Vocal expression of affect. In *The new handbook of methods in nonverbal behavior research*: 65–135. Harrigan, J., Rosenthal, R. & Scherer, K. (Eds). Oxford: Oxford University Press.
- Keltner, D. & Lerner, J.S. (2010). Emotion. In *Handbook of social psychology*: 317–352. Gilbert, D.T., Fiske, S.T. & Lindzey, G. (Eds). New York: Wiley.
- Knutson, B., Burgdorf, J. & Panksepp, J. (2002). Ultrasonic vocalizations as indices of affective states in rats. *Psychol. Bull.* **128**, 961–977.
- Langford, D.J., Bailey, A.L., Chanda, M.L., Clarke, S.E., Drummond, T.E., Echols, S., Ingrao J., Klassen-Ross T., Lacroix-Fralish M.L., Matsumiya L., Sorge R.E., Sotocinal S.G., Tabaka J.M., Wong D., van den Maagdenberg A.M., Ferrari M.D., Craig K.D. & Mogil J.S. (2010). Coding of facial expressions of pain in the laboratory mouse. *Nat. Methods* **7**, 447–449.
- Laukkanen, A.-M., Vilkmann, E., Alku, P. & Oksanen, H. (1997). On the perception of emotions in speech: the role of voice quality. *Logoped. Phoniatr. Vocol.* **22**, 157–168.
- Li, X., Tao, J., Johnson, M., Soltis, J., Savage, A., Leong, K. & Newman, J. (2007). Stress and emotion classification using jitter and shimmer features. Proceedings of the Acoustics, Speech and Signal Processing, Honolulu, HI.
- Lord, K., Feinstein, M. & Coppinger, R. (2009). Barking and mobbing. *Behav. Proc.* **81**, 358–368.
- Macedonia, J.M. (1993). The vocal repertoire of the ringtailed lemur (*Lemur catta*). *Folia Primatol.* **61**, 186–217.
- Manser, M.B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proc. R. Soc. Lond. Ser. B* **268**, 2315–2324.
- Manser, M.B. (2010). The generation of functionally referential and motivational vocal signals in mammals. In *Handbook of mammalian vocalization – an integrative neuroscience approach*: 477–486. Brudzynski, S.M. (Ed.). London: Academic Press.
- Manser, M.B., Seyfarth, R.M. & Cheney, D.L. (2002). Suricate alarm calls signal predator class and urgency. *Trends Cogn. Sci.* **6**, 55–57.
- Manteuffel, G., Puppe, B. & Schön, P.-C. (2004). Vocalization of farm animals as a measure of welfare. *Appl. Anim. Behav. Sci.* **88**, 163–182.
- Marchant, J.N., Whittaker, X. & Broom, D.M. (2001). Vocalisations of the adult female domestic pig during a standard human approach test and their relationships with behavioural and heart rate measures. *Appl. Anim. Behav. Sci.* **72**, 23–39.
- McComb, K., Taylor, A.M., Wilson, C. & Charlton, B.D. (2009). The cry embedded within the purr. *Curr. Biol.* **19**, R507–R508.
- McElligott, A.G., Birrer, M. & Vannoni, E. (2006). Retraction of the mobile descended larynx during groaning enables fallow bucks (*Dama dama*) to lower their formant frequencies. *J. Zool.* **270**, 340–345.

- Meise, K., Keller, C., Cowlshaw, G. & Fischer, J. (2011). Sources of acoustic variation: implications for production specificity and call categorization in chacma baboon (*Papio ursinus*) grunts. *J. Acoust. Soc. Am.* **129**, 1631–1641.
- Mendl, M., Burman, O.H.P. & Paul, E.S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proc. R. Soc. Lond. Ser. B* **277**, 2895–2904.
- Monticelli, P.F., Tokumaru, R.S. & Ades, C. (2004). Isolation induced changes in guinea pig *Cavia porcellus* pup distress whistles. *An. Acad. Bras. Cienc.* **76**, 368–372.
- Morton, E.S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* **111**, 855–869.
- Murray, I.R. & Arnott, J.L. (1993). Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. *J. Acoust. Soc. Am.* **93**, 1097–1108.
- Norcross, J.L. & Newman, J.D. (1999). Effects of separation and novelty on distress vocalizations and cortisol in the common marmoset (*Callithrix jacchus*). *Am. J. Primatol.* **47**, 209–222.
- Norcross, J.L., Newman, J.D. & Cofrancesco, L.M. (1999). Context and sex differences exist in the acoustic structure of phee calls by newly-paired common marmosets (*Callithrix jacchus*). *Am. J. Primatol.* **49**, 165–181.
- Owren, M.J., Seyfarth, R.M. & Cheney, D.L. (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.
- Panksepp, J. (2005). Affective consciousness: core emotional feelings in animals and humans. *Conscious. Cogn.* **14**, 30–80.
- Panksepp, J. (2009). Emotional causes and consequences of social-affective vocalization. In *Handbook of mammalian vocalization – an integrative neuroscience approach*: 201–208. Brudzynski, S.M. (Ed.). London: Academic Press.
- Panksepp, J. (2011). The basic emotional circuits of mammalian brains: do animals have affective lives? *Neurosci. Biobehav. Rev.* **35**, 1791–1804.
- Panksepp, J. & Burgdorf, J. (2000). 50-kHz chirping (laughter?) in response to conditioned and unconditioned tickle-induced reward in rats: effects of social housing and genetic variables. *Behav. Brain Res.* **115**, 25–38.
- Panksepp, J. & Burgdorf, J. (2003). ‘Laughing’ rats and the evolutionary antecedents of human joy? *Physiol. Behav.* **79**, 533–547.
- Patel, S., Scherer, K.R., Björkner, E. & Sundberg, J. (2011). Mapping emotions into acoustic space: the role of voice production. *Biol. Psychol.* **87**, 93–98.
- Paul, E., Harding, E.J. & Mendl, M. (2005). Measuring emotional processes in animals: the utility of a cognitive approach. *Neurosci. Biobehav. Rev.* **29**, 469–491.
- Pell, M., Paulmann, S., Dara, S., Alasser, A. & Kotz, S. (2008). Similarities in the acoustic expression of emotions in English, German, Hindi, and Arabic. *J. Acoust. Soc. Am.* **124**, 2496.
- Peters, G. (2002). Purring and similar vocalizations in mammals. *Mamm. Rev.* **32**, 245–271.
- Pond, R.L., Darre, M.J., Scheifele, P.M. & Browning, D.G. (2010). Characterization of equine vocalization. *J. Vet. Behav.* **5**, 7–12.
- Pongrácz, P., Csaba, M. & Miklósi, A. (2006). Acoustic parameters of dog barks carry emotional information for humans. *Appl. Anim. Behav. Sci.* **100**, 228–240.
- Puppe, B., Schön, P.-C., Tuchscherer, A. & Manteuffel, G. (2005). Castration-induced vocalisation in domestic piglets, *Sus scrofa*: complex and specific alterations of the vocal quality. *Appl. Anim. Behav. Sci.* **95**, 67–78.
- Reby, D. & McComb, K. (2003). Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Anim. Behav.* **65**, 519–530.
- Reefmann, N., Bütikofer Kaszás, F., Wechsler, B. & Gygas, L. (2009a). Ear and tail postures as indicators of emotional valence in sheep. *Appl. Anim. Behav. Sci.* **118**, 199–207.
- Reefmann, N., Wechsler, B. & Gygas, L. (2009b). Behavioural and physiological assessment of positive and negative emotion in sheep. *Anim. Behav.* **78**, 651–659.
- Rendall, D. (2003). 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., Seyfarth, R.M., Cheney, D.L. & Owren, M.J. (1999). The meaning and function of grunt variants in baboons. *Anim. Behav.* **57**, 583–592.
- Robbins, R.L. & McCreery, E.K. (2003). African wild dog pup vocalizations with special reference to Mortons model. *Behaviour* **140**, 333–351.
- Roessler, R. & Lester, J.W. (1976). Voice predicts affect during psychotherapy. *J. Nerv. Ment. Dis.* **163**, 166–176.
- Ross, M.D., Owren, M.J. & Zimmermann, E. (2009). Reconstructing the evolution of laughter in great apes and humans. *Curr. Biol.* **19**, 1106–1111.
- Sauter, D.A., Eisner, F., Ekman, P. & Scott, S.K. (2010). Cross-cultural recognition of basic emotions through non-verbal emotional vocalizations. *Proc. Natl. Acad. Sci. USA* **107**, 2408–2412.
- Schehka, S., Esser, K.-H. & Zimmermann, E. (2007). Acoustical expression of arousal in conflict situations in tree shrews (*Tupaia belangeri*). *J. Comp. Physiol. A* **193**, 845–852.
- Schehka, S. & Zimmermann, E. (2009). Acoustic features to arousal and identity in disturbance calls of tree shrews (*Tupaia belangeri*). *Behav. Brain Res.* **203**, 223–231.
- Scheiner, E. & Fisher, J. (2011). Emotion expression – the evolutionary heritage in the human voice. In *Interdisciplinary anthropology: the continuing evolution of man*: 105–130. Welsch, W., Singer, W. & Wunder, A. (Eds). Heidelberg & New York: Springer.

- Scherer, K.R. (1986). Vocal affect expression: a review and a model for future research. *Psychol. Bull.* **99**, 143–165.
- Scherer, K.R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Commun.* **40**, 227–256.
- Scherer, K.R., Banse, R. & Wallbott, H.G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *J. Cross Cult. Psychol.* **32**, 76–92.
- Scherer, K.R., Feldstein, S., Bond, R.N. & Rosenthal, R. (1985). Vocal cues to deception: a comparative channel approach. *J. Psycholinguist. Res.* **14**, 409–425.
- Scherer, K.R. & Kappas, A. (1988). Primate vocal expression of affective state. In *Primate vocal communication*: 171–194. Todt, D., Goedeke, P. & Newman, E. (Eds). Heidelberg: Springer.
- Scherer, K.R., Ladd, D.R. & Silverman, K.E.A. (1984). Vocal cues to speaker affect: testing two models. *J. Acoust. Soc. Am.* **76**, 1346–1356.
- Scheumann, M., Zimmermann, E. & Deichsel, G. (2007). Context-specific calls signal infants' needs in a strepsirrhine primate, the gray mouse lemur (*Microcebus murinus*). *Dev. Psychobiol.* **49**, 708–718.
- Schön, P.-C., Puppe, B. & Manteuffel, G. (2004). Automated recording of stress vocalization as a tool to document impaired welfare in pigs. *Anim. Welf.* **13**, 105–110.
- Schrader, L. & Todt, D. (1993). Contact call parameters covary with social context in common marmosets *Callithrix j. jacchus*. *Anim. Behav.* **46**, 1026–1028.
- Schrader, L. & Todt, D. (1998). Vocal quality is correlated with levels of stress hormones in domestic pigs. *Ethology* **104**, 859–876.
- Sèbe, F., Poindron, P., Andanson, S., Chandeze, H., Delval, E., Després, G., Toporenko, G., Bickell, S., Blache, D. & Boissy, A. (2012). Bio-acoustic analyses to assess emotion in animals: acoustic patterns are linked to behavioural, cardiac and hormonal responses of ewes to the separation from their lambs. Proceedings of the International Bio-acoustics Council meeting, Lisbon, Portugal. *Bioacoustics* **51**, 54.
- Seyfarth, R.M. & Cheney, D.L. (2003). Meaning and emotion in animal vocalizations. *Ann. N. Y. Acad. Sci.* **1000**, 32–55.
- Shair, H.N., Brunelli, S.A., Masmela, J.R., Boone, E. & Hofer, M.A. (2003). Social, thermal, and temporal influences on isolation-induced and maternally potentiated ultrasonic vocalizations of rat pups. *Dev. Psychobiol.* **42**, 206–222.
- Siebert, E.R. & Parr, L.A. (2003). A structural and contextual analysis of chimpanzee screams. *Ann. N. Y. Acad. Sci.* **1000**, 104–109.
- Siebert, K., Langbein, J., Schön, P.-C., Tuchscherer, A. & Puppe, B. (2011). Degree of social isolation affects behavioural and vocal response patterns in dwarf goats (*Capra hircus*). *Appl. Anim. Behav. Sci.* **131**, 53–62.
- Siegel, A., Bhatt, S., Bhatt, R. & Zalcman, S.S. (2010). Limbic, hypothalamic and periaqueductal gray circuitry and mechanisms controlling rage and vocalization in the cat. In *Handbook of mammalian vocalization – an integrative neuroscience approach*: 243–253. Brudzynski, S.M. (Ed.). London: Academic Press.
- Slocumbe, K.E. & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proc. Natl. Acad. Sci. USA* **104**, 17228–17233.
- Soltis, J., Blowers, T.E. & Savage, A. (2011). Measuring positive and negative affect in the voiced sounds of African elephants (*Loxodonta africana*). *J. Acoust. Soc. Am.* **129**, 1059–1066.
- Soltis, J., Leighty, K.A., Wesolek, C.M. & Savage, A. (2009). The expression of affect in African elephant (*Loxodonta africana*) rumble vocalizations. *J. Comp. Psychol.* **123**, 222–225.
- Soltis, J., Leong, K. & Savage, A. (2005). African elephant vocal communication II: rumble variation reflects the individual identity and emotional state of callers. *Anim. Behav.* **70**, 589–599.
- Stoeger, A.S., Charlton, B.D., Kratochvil, H. & Fitch, W.T. (2011). Vocal cues indicate level of arousal in infant African elephant roars. *J. Acoust. Soc. Am.* **130**, 1700–1710.
- Sugiura, H. (2007). Effects of proximity and behavioral context on acoustic variation in the coo calls of Japanese macaques. *Am. J. Primatol.* **69**, 1412–1424.
- Szameitat, D.P., Darwin, C.J., Szameitat, A.J., Wildgruber, D. & Alter, K. (2011). Formant characteristics of human laughter. *J. Voice* **25**, 32–37.
- Tate, A.J., Fischer, H., Leigh, A. & Kendrick, K.M. (2006). Behavioural and neurophysiological evidence for face identity and face emotion processing in animals. *Philos. Trans. R. Soc. Lond. B* **361**, 2155–2172.
- Taylor, A.M. & Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research. *J. Zool.* **280**, 221–236.
- Taylor, A.M., Reby, D. & McComb, K. (2009). Context-related variation in the vocal growling behaviour of the domestic dog (*Canis familiaris*). *Ethology* **115**, 905–915.
- Theis, K.R., Greene, K.M., Benson-Amram, S.R. & Holkamp, K.E. (2007). Sources of variation in the long-distance vocalizations of spotted hyenas. *Behaviour* **144**, 557–584.
- Thomas, T.J., Weary, D.M. & Appleby, M.C. (2001). Newborn and 5-week-old calves vocalize in response to milk deprivation. *Appl. Anim. Behav. Sci.* **74**, 165–173.
- Titze, I.R. (1994). *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.
- Tolkmitt, F.J. & Scherer, K.R. (1986). Effects of experimentally induced stress on vocal parameters. *J. Exp. Psychol. Hum. Percept. Perform.* **12**, 302–313.
- Waaramaa, T., Alku, P. & Laukkanen, A.-M. (2006). The role of F3 in the vocal expression of emotions. *Logoped. Phoniatr. Vocol.* **31**, 153–156.
- Waaramaa, T., Laukkanen, A.-M., Airas, M. & Alku, P. (2010). Perception of emotional valences and activity levels

- from vowel segments of continuous speech. *J. Voice* **24**, 30–38.
- Watts, J.M. & Stookey, J.M. (1999). Effects of restraint and branding on rates and acoustic parameters of vocalization in beef cattle. *Appl. Anim. Behav. Sci.* **62**, 125–135.
- Watts, J.M. & Stookey, J.M. (2000). Vocal behaviour in cattle: the animal's commentary on its biological processes and welfare. *Appl. Anim. Behav. Sci.* **67**, 15–33.
- Weary, D.M., Braithwaite, L.A. & Fraser, D. (1998). Vocal response to pain in piglets. *Appl. Anim. Behav. Sci.* **56**, 161–172.
- Weary, D.M. & Fraser, D. (1995a). Calling by domestic piglets: reliable signals of need? *Anim. Behav.* **50**, 1047–1055.
- Weary, D.M. & Fraser, D. (1995b). Signalling need: costly signals and animal welfare assessment. *Appl. Anim. Behav. Sci.* **44**, 159–169.
- Weary, D.M., Ross, S. & Fraser, D. (1997). Vocalizations by isolated piglets: a reliable indicator of piglet need directed towards the sow. *Appl. Anim. Behav. Sci.* **53**, 249–257.
- Wich, S.A., Swartz, K.B., Hardus, M.E., Lameira, A.R., Stromberg, E. & Shumaker, R. (2009). A case of spontaneous acquisition of a human sound by an orangutan. *Primates* **50**, 56–64.
- Williams, C.E. & Stevens, K.N. (1972). Emotions and speech: some acoustical correlates. *J. Acoust. Soc. Am.* **52**, 1238–1250.
- 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.
- Yamaguchi, C., Izumi, A. & Nakamura, K. (2010). Time course of vocal modulation during isolation in common marmosets (*Callithrix jacchus*). *Am. J. Primatol.* **72**, 681–688.
- Yeon, S.C., Kim, Y.K., Park, S.J., Lee, S.S., Lee, S.Y., Suh, E.H., Houpt, K.A., Chang H.H., Lee H.C., Yang B.G. & Lee H.J. (2011). Differences between vocalization evoked by social stimuli in feral cats and house cats. *Behav. Proc.* **87**, 183–189.
- Yin, S. & McCowan, B. (2004). Barking in domestic dogs: context specificity and individual identification. *Anim. Behav.* **68**, 343–355.
- Zahavi, A. (1982). The pattern of vocal signals and the information they convey. *Behaviour* **80**, 1–8.
- Zei Pollermann, B. (2008). Autonomic determinants of vocal expression of emotions. In *Emotions in the human voice*: 1–16. Izdebski, K. (Ed.). San Diego, CA: Plural Publishing.
- Zei Pollermann, B. & Archinard, M. (2002). Acoustic patterns of emotions. In *Improvements in speech synthesis*: 237–245. Keller, E., Bailly, G., Monaghan, A., Terken, J. & Huckvale, M. (Eds). Chichester: John Wiley & Sons.
- Zimmermann, E. (2009). Vocal expression of emotion in a nocturnal prosimian primate group, mouse lemurs. In *Handbook of mammalian vocalization – an integrative neuroscience approach*: 215–225. Brudzynski, S.M. (Ed.). London: Academic Press.