Author for correspondence:
Tamás Faragó
e-mail: mustela.nivalis@gmail.com

Royal Society Publishing

Animal behaviour

Humans rely on the same rules to assess emotional valence and intensity in conspecific and dog vocalizations

Tamás Faragó[1], Attila Andics[1], Viktor Devecseri[1], Anna Kis[2,3], Márta Gácsi[1] and Ádám Miklósi[1,2]

[1]MTA-ELTE Comparative Ethology Research Group, Pázmány Péter sétány 1/C, Budapest 1117, Hungary,
[2]Department of Ethology, Eötvös University, Pázmány Péter sétány 1/C, Budapest 1117, Hungary
[3]Research Centre for Natural Sciences, Hungarian Academy of Sciences, Victor Hugo utca 18-22, Budapest 1132, Hungary

Humans excel at assessing conspecific emotional valence and intensity, based solely on non-verbal vocal bursts that are also common in other mammals. It is not known, however, whether human listeners rely on similar acoustic cues to assess emotional content in conspecific and heterospecific vocalizations, and which acoustical parameters affect their performance. Here, for the first time, we directly compared the emotional valence and intensity perception of dog and human non-verbal vocalizations. We revealed similar relationships between acoustic features and emotional valence and intensity ratings of human and dog vocalizations: those with shorter call lengths were rated as more positive, whereas those with a higher pitch were rated as more intense. Our findings demonstrate that humans rate conspecific emotional vocalizations along basic acoustic rules, and that they apply similar rules when processing dog vocal expressions. This suggests that humans may utilize similar mental mechanisms for recognizing human and heterospecific vocal emotions.

1. Introduction

Emotions are an organism's specialized mental states, shaped by natural selection, enabling them to increase fitness in certain contexts by facilitating adaptive physiological, cognitive and behavioural responses [1]. Non-linguistic vocal emotional expressions are ancient, evolutionarily conservative, easily recognized by humans [2] and less affected by cultural differences than prosody or linguistic emotional expressions [3]. Most emotional vocalizations consist of calls that are acoustically highly similar in both humans and other species [4]. These calls, as the smallest meaningful units, are the building blocks of vocal emotion expressions and their acoustic properties affect how listeners perceive their emotional content [5].

According to the 'pre-human origin' hypothesis of affective prosody, the acoustic cues of emotions in human vocalizations are innate and have strong evolutionary roots [6]. Furthermore, according to the source-filter framework, the basic mechanisms of sound production are the same among human and terrestrial vertebrates [7], suggesting that similar vocal parameters may carry information for the listeners about the caller's inner state [8]. We can therefore hypothesize that similar basic rules support vocal emotion recognition both within and across species. However, we have little information about how the wide variety of possible emotional states are encoded and perceived in vocalizations, and whether terrestrial vertebrates use these parameters or follow other rules when processing emotional sounds.

Dogs, owing to their special status in human society [9] and their numerous vocalization types used in various social contexts [10], can provide an excellent insight into this question. Recent studies showed that the acoustics of dog barks affects humans' inner state assessment following the Morton's motivation–structural (MS) rules [11]: low-pitched barks with short inter-bark intervals were rated as aggressive, whereas high-pitched ones with long intervals were considered playful and happy. However, it is not clear yet whether the same principle holds across their diverse vocal repertoire. More importantly, based on the analogous vocal production mechanisms and the pristine nature of the human non-verbal vocal emotional expressions, we predict that humans use similar features of human and dog vocalizations to assess the signaller's inner state. Thus, our aim was to compare which basic acoustical properties of dog and human vocalizations affect how human listeners assess their emotional content.

There are two main approaches to studying emotions. The framework of discrete emotions is rooted in studies of human facial expressions and claims to focus on 'pure' emotions. By contrast, dimensional models aim to account for gradedness found in studies of subjective experiences of emotions and suggest that inner states can be effectively modelled as coordinates of a two- or three-dimensional space [12]. Following Mendl's suggestion, we adapted Russell's widely used dimensional model [13] and asked listeners to rate human and dog vocalizations along two axes: (i) emotional valence, ranging from negative to positive, and (ii) emotional intensity (we use this term as a synonym of emotional arousal as in [6]), ranging from non-intense to intense. To explore how specific acoustic cues affect the ratings of human and dog vocalizations, we measured four basic parameters for each vocal stimulus. Fundamental frequency (F_0) and tonality (harmonic-to-noise ratio: HNR) are used as inner state indicators in the MS rules, and, as source-related parameters, they are potentially affected by arousal and emotional quality [7]. In addition, spectral centre-of-gravity (CG) of vocalizations had been found to affect the perception of valence and arousal in human non-verbal vocalizations [5]. Finally, the average call length (CL) within a sound sample is also a commonly measured temporal parameter linked to the emotional state of the signaller [6,8].

2. Material and methods

Our subjects were Hungarian volunteers, recruited via the Internet and through personal requests (six males, 33 females, age: 31 ± 9 years; electronic supplementary material, table S3). We compiled a pool of 100–100 non-verbal vocalizations of dogs and humans from diverse social contexts and various sound types (for details, see the electronic supplementary material). Acoustical measurements were carried out with a semiautomatic Praat script. First, each basic vocal unit within a sound sample was marked, to be considered later on as an individual call (see electronic supplementary material, figure S1, and a similar approach in [4]). We measured CL, F_0, HNR and CG in each call. Then, these call-by-call measurements were averaged within each sound sample to characterize each sample by one value of the given parameter (standard deviation across all calls was two to five times greater than the average within-sample standard deviation, for all acoustic variables).

A novel online-based survey (http://www.inflab.bme.hu/~viktor/sr_demo/) was developed to assess how humans perceive the emotional content of vocalizations. Instead of using
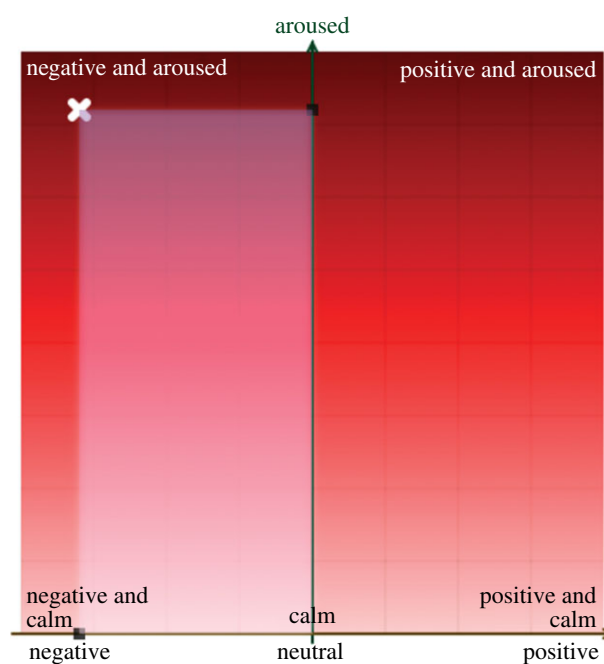

Figure 1. Two-dimensional response plane. *x*- and *y*-axes represent emotional valence and emotional intensity, respectively. The system projects the cursor's position to both axes (opaque area) to visually help the subjects to rate stimuli along both dimensions at the same time. The white X shows where the subject clicked to rate the actual sound. (Online version in colour.)

independent basic emotion scales, we applied a slightly modified version of Russell's two-dimensional model (figure 1). Subjects rated the emotional valence and intensity of the sounds by clicking on one point of a coordinate system. The system registered the two coordinates (valence: −50 to 50; intensity: 0–100) and the reaction time. After three practice trials, all 200 stimuli were presented randomly. Every sample was played once for each subject, except for 5–5 selected and randomly repeated dog and human samples used to test the reliability of subjects' responses (see the electronic supplementary material). We also added two breaks, unrestricted in length, after the 70th and the 141st sound. We analysed the data of those ($N = 39$) subjects who completed the survey for all sound samples.

To reveal the effects of acoustic parameters on the responses, multivariate linear regressions were applied. For this, we averaged the valence and intensity ratings within each sample. We used a backward elimination method to find the parameters that affected the ratings most (one dog sample was excluded owing to its extremely high fundamental frequency—3500 Hz).

3. Results

The regression models showed significant relationships between emotional ratings and acoustic measures for both dog and human vocalizations. We found that valence ratings were affected by CL in both dog and human samples: the shorter the calls were within a sound sample, the more positively the sample was rated. For human sounds, lower CG values corresponded with more positive valence scores. The intensity scale was also affected by the measured acoustical parameters in both human and dog sounds. Partial regressions showed that the intensity was influenced by F_0: higher pitched samples were rated as more intense in both species' vocalizations. We also found species-specific effects where the intensity ratings of dog samples were affected by the change of the other three
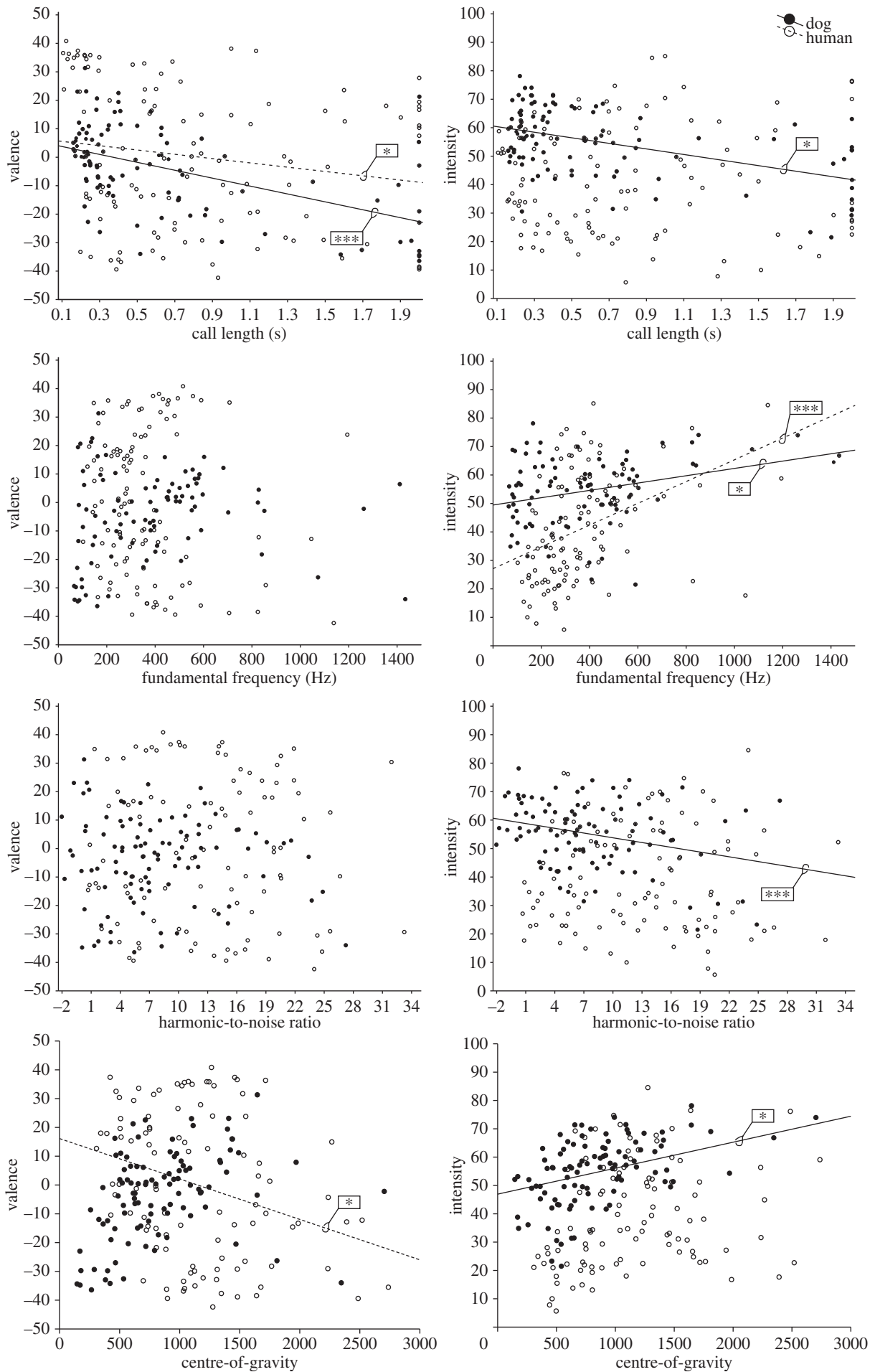
Figure 2. The linear relationships between acoustic parameters and emotional scales. Filled circles represent dog vocalizations, and open circles represent human vocalizations. The asterisks show significant relationship between the measures (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

Table 1. The results of the multivariate linear regressions. The table shows only the data of models obtained by backward elimination (model lines), and the partial regressions of the remaining parameters. The grey cells show the eliminated parameters. Std β, standardized beta value; CL, call length; F_0, fundamental frequency; HNR, harmonic-to-noise ratio; CG, spectral centre-of-gravity.

acoustic parameters		valence			intensity		
dog	**model**	R^2	F	p	R^2	F	p
		0.292	37.564	0.00	0.436	17.01	0.00
		std β	t	p	std β	t	p
	partial regressions						
	CL	−0.541	−6.129	0.00	−0.215	−2.214	0.03
	F_0				0.320	2.423	0.02
	HNR				−0.453	−4.345	0.00
	CG				0.235	2.088	0.04
human	**model**	R^2	F	p	R^2	F	p
		0.119	6.572	0.00	0.179	21.387	0.00
		std β	t	p	std β	t	p
	partial regressions						
	CL	−0.188	−1.969	0.05			
	F_0				0.423	4.625	0.00
	HNR						
	CG	−0.282	−2.951	0.00			

acoustical parameters: longer and more tonal dog samples were rated as less intense, while higher CG was related to higher intensity ratings (figure 2 and for statistical details, see table 1).

4. Discussion

This study is the first to directly compare how humans perceive human and dog emotional vocalizations. We show that humans use similar acoustical parameters to attribute emotional valence and intensity to both human and dog vocalizations.

Our results support the pre-human origin hypothesis of affective prosody [6] and are indicative of similar mechanisms underlying the processing of human and dog vocal emotion expressions. Evolutionarily ancient systems could possibly be used for processing the emotional load of non-linguistic human and non-human vocalizations. Alternatively, humans may judge the emotional states of non-human animal sounds on the basis of perceived acoustic similarity to their own vocalizations.

Our results are in agreement with previous studies aiming to assess the acoustic rules underlying the processing of different vocalizations. However, we also reveal novel and previously unexplored relationships. Pongrácz *et al.* [11] found that, in case of dog barks, deeper pitch and fast pulsing can be linked to higher aggression, while low pulsing and higher pitch to positive valence, and higher tonality to higher despair ratings. By contrast, our results show that, with regards to dog vocalizations, long, high-pitched and tonal sounds can be linked to fearful inner states (high intensity, negative valence), long, low-pitched, noisy sounds to aggressiveness (lower intensity, still negative valence) and short, pulsing sounds independent of their pitch and tonality are connected to positive inner states. As barks are highly specialized vocalizations of dogs formed by domestication

[14], no general rule can be drawn based on their acoustical structure. In our study, the high diversity of calls showed clear parallels with assessments of human emotional valence and intensity. Sauter *et al.* [5] reported similar effects of fundamental frequency in intensity and spectral-CG in valence ratings, while, in contrast with our results, they found CL affecting the intensity ratings negatively and spectral-CG positively. These differences may be due to the different composition of the presented stimuli. While Sauter's recordings of ten acted emotions originated from four adult vocalizers, our sample (although still not fully representative) covered a wider range of call types and vocalizers, resulting in higher acoustic variance and revealing different connections. Besides the basic parameters investigated here, several others may play a role in valence perception (for a review, see [8]). While our within-sample averaging approach was insensitive to the within-sample dynamics of acoustic parameters across consecutive calls, such dynamic changes may also convey relevant information about the inner state of the signaller. More studies are needed to determine whether acoustic similarities between human and dog vocalizations also reflect functional similarity in their emotional states.

To conclude, our results provide the first evidence of the use of the same basic acoustic rules in humans for the assessment of emotional valence and intensity in both human and dog vocalizations. Further comparative studies using vocalizations from a wide variety of species may reveal the existence of a common mammalian basis for emotion communication, as suggested by our results.

References

1. Nesse RM. 1990 Evolutionary explanations of emotions. *Hum. Nat.* **1**, 261–289. (doi:10.1007/BF02733986)

2. Simon-Thomas ER, Keltner DJ, Sauter DA, Sinicropi-Yao L, Abramson A. 2009 The voice conveys specific emotions: evidence from vocal burst displays. *Emotion* **9**, 838–846. (doi:10.1037/a0017810)

3. Sauter DA, Eisner F, Ekman P, Scott SK. 2010 Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl Acad. Sci. USA* **107**, 2408–2412. (doi:10.1073/pnas.0908239106)

4. Bachorowski J-A, Smoski MJ, Owren MJ. 2001 The acoustic features of human laughter. *J. Acoust. Soc. Am.* **110**, 1581. (doi:10.1121/1.1391244)

5. Sauter DA, Eisner F, Calder AJ, Scott SK. 2010 Perceptual cues in nonverbal vocal expressions of emotion. *Q. J. Exp. Psychol.* **63**, 2251–2272. (doi:10.1080/17470211003721642)

6. Zimmermann E, Leliveld L, Schehka S. 2013 Toward the evolutionary roots of affective prosody in human acoustic communication: a comparative approach to mammalian voices. In *Evolution of emotional communication: from sounds in nonhuman mammals to speech and music in man.* (eds E Altenmüller, S Schmidt, E Zimmermann), pp. 116–132. Oxford, UK: Oxford University Press.

7. Taylor AM, Reby D. 2010 The contribution of source-filter theory to mammal vocal communication research. *J. Zool.* **280**, 221–236. (doi:10.1111/j.1469-7998.2009.00661.x)

8. Briefer EF. 2012 Vocal expression of emotions in mammals: mechanisms of production and evidence. *J. Zool.* **288**, 1–20. (doi:10.1111/j.1469-7998.2012.00920.x)

9. Topál J, Miklósi Á, Gácsi M, Dóka A, Pongrácz P, Kubinyi E, Virányi Z, Csányi V. 2009 The dog as a model for understanding human social behavior. *Adv. Study Behav.* **39**, 71–116. (doi:10.1016/S0065-3454(09)39003-8)

10. Cohen JA, Fox MW. 1976 Vocalizations in wild canids and possible effects of domestication. *Behav. Process.* **1**, 77–92. (doi:10.1016/0376-6357(76)90008-5)

11. Pongrácz P, Molnár C, Miklósi Á. 2006 Acoustic parameters of dog barks carry emotional information for humans. *Appl. Anim. Behav. Sci.* **100**, 228–240. (doi:10.1016/j.applanim.2005.12.004)

12. Mendl MT, Burman OHP, Paul ES. 2010 An integrative and functional framework for the study of animal emotion and mood. *Proc. R. Soc. B* **277**, 2895–2904. (doi:10.1098/rspb.2010.0303)

13. Russell JA. 1980 A circumplex model of affect. *J. Pers. Soc. Psychol.* **39**, 1161–1178. (doi:10.1037/h0077714)

14. Pongrácz P, Molnár C, Miklósi Á. 2010 Barking in family dogs: an ethological approach. *Vet. J.* **183**, 141–147. (doi:10.1016/j.tvjl.2008.12.010)