

## **Voice-sensitive regions in the dog and human brain are revealed by comparative fMRI**

Attila Andics<sup>1,\*</sup>, Márta Gácsi<sup>1</sup>, Tamás Faragó<sup>1</sup>, Anna Kis<sup>2,3</sup>, Ádám Miklósi<sup>1,2</sup>

<sup>1</sup>MTA-ELTE Comparative Ethology Research Group, Budapest, Hungary. H-1117 Budapest,  
Pázmány Péter sétány 1/C.

<sup>2</sup>Department of Ethology, Eötvös Loránd University, Budapest, Hungary. H-1117 Budapest, Pázmány  
Péter sétány 1/C.

<sup>3</sup>Research Centre for Natural Sciences, Institute of Cognitive Neuroscience and Psychology,  
Hungarian Academy of Sciences, Budapest, Hungary. H-1117 Budapest, Magyar Tudósok körútja 2.

\*Correspondence:

E-mail: attila.andics@gmail.com

Phone: +36 1 381 21 79

Fax: +36 1 381 21 80

Running title: **Voice-sensitive regions in the dog and human brain**

## **Highlights**

- The first comparative neuroimaging study of a non-primate species and humans
- Functional analogies in non-primary auditory cortex between dogs and humans
- Voice areas preferring conspecific vocalizations were evidenced in the dog brain
- Brain sensitivity to acoustic cues of vocal emotional information in both species

## **Summary**

During the approximately 18-32 thousand years of domestication [1], dogs and humans have shared a similar social environment [2]. Dog and human vocalizations are thus familiar and relevant to both species [3], although they belong to evolutionarily distant taxa, as their lineages split approximately 90-100 million years ago [4]. In this first comparative neuroimaging study of a non-primate and a primate species, we made use of this special combination of shared environment and evolutionary distance. We presented dogs and humans with the same set of vocal and non-vocal stimuli to search for functionally analogous voice-sensitive cortical regions. We demonstrate that voice areas exist in dogs, and that they show a similar pattern to anterior temporal voice areas in humans. Our findings also reveal that sensitivity to vocal emotional valence cues engages similarly located non-primary auditory regions in dogs and humans. Although parallel evolution cannot be excluded, our findings suggest that voice areas may have a more ancient evolutionary origin than previously known.

## **Results and Discussion**

An important social function of the auditory system is to process the vocalizer's identity and emotional state. Non-primary auditory brain regions preferring conspecific vocalizations were found in both humans [5, 6] and non-human primates, suggesting that 'voice areas' evolved at least 30 million years ago [7–10]. In humans, auditory regions sensitive to vocal emotional cues have also been identified [11–14]. Research has also indicated that vocal emotional valence is conveyed via similar acoustic rules across species [15], including human [16] and non-human animals [17].

Behavioural field research has revealed that the efficient processing of conspecificity and emotional information in vocalizations is important in both primate [18–20] and non-primate species [21–24]. Indeed, both the acoustic recognition of conspecifics and their emotional state is fundamental for making decisions in behaviour contexts like mate choice, territory disputes or hierarchy-related challenges [25]. Nevertheless, little is known about the underlying neural mechanisms of vocalization processing in non-primates.

To reveal possible functional analogies between human and non-primate auditory brain regions, this study describes a comparative investigation of dogs and humans. We investigated (1) whether in dogs, similarly to humans, certain auditory regions ('voice areas') would respond stronger to conspecific vocalizations than to either heterospecific vocalizations or non-vocal sounds, and (2) whether dogs are similar to humans in the cortical processing of emotional cues in vocal signals.

To address these questions, we used a non-invasive functional magnetic resonance imaging (fMRI) procedure with awake dogs (N=11) and humans (N=22). We built on our group's first small sample-sized attempts of awake dog fMRI [26], providing a procedure different from others' [27]. All participants were unrestrained and instructed to lay motionless in an fMRI

scanner for three 6-min runs (see Figure 1 and Figure S1). Dogs and humans listened to an identical set of stimuli, which included three sound types: human vocalizations, dog vocalizations, non-vocal environmental sounds, and a silent baseline. Vocal stimuli ranged parametrically in emotional valence from highly negative to highly positive, as rated by an independent set of human listeners [16]. Neural sensitivity to conspecificity and emotional valence (and related acoustic cues) was evaluated similarly for the two species using random effects group analyses (see Supplemental Experimental Procedures).

Auditory regions were defined functionally, using the all sounds vs silence contrast (Figure 2a,b). Similar sound-sensitive brain regions were identified in dogs and humans, including regions of the auditory cortex and subcortical regions (Table S1). Consistent with lesion studies on dog auditory processing regions [28, 29], cortical sound-sensitivity in dogs was localized in perisylvian regions, including the Sylvian gyri (SG) along the Sylvian fissure (SF), the ectosylvian gyri (ESG) along the ectosylvian sulcus (ESS), and extending dorsally to the suprasylvian sulcus (SSS). Human cortical auditory activity was found along the superior temporal sulcus (STS) and in the inferior frontal cortex (IFC). In both species, auditory activity extended ventrally towards the temporal pole (TP), i.e. the most basal part of the caudal SG in dogs, and the anterior tip of the temporal lobe in humans. A subcortical sound-sensitive region, with a peak in the first-order auditory thalamus, the medial geniculate body (MGB, [30, 31]), including the caudal colliculus and extending towards the cerebellum was also identified in both species. The search space of all following analyses was defined by these functionally localized auditory regions: their total size was 12 cm<sup>3</sup> (1441 voxels) for dogs and 95 cm<sup>3</sup> (11849 voxels) for humans.

A first qualitative comparison of parameter estimates for each sound type indicates an important difference between dog and human auditory regions. Dogs have subregions where parameter estimates were maximal for dog vocalizations (39% of all auditory voxels), but also

subregions with maximal response to human vocal (13%) or non-vocal (48%) sounds. In contrast, almost all human auditory regions were maximal for human vocalizations (87%). Maximal response for dog vocalizations (10%) was found in the subcortical MGB, and almost no subregions were found where the response was maximal for non-vocal sounds (3%) (Figure 2c).

To identify voice areas, i.e. auditory regions responding preferentially to conspecific vocalizations compared to either heterospecific or non-vocal sounds, brain responses to each sound type were compared in random effects conjunction analyses (Figure 3, Table S2 and Supplemental Experimental Procedures). In dogs, we identified a ventral auditory region (cSG), close to the TP bilaterally and a left dorsal auditory region (mESG) that responded stronger to dog than to either human or non-vocal sounds. None of these regions responded stronger to human than to non-vocal sounds. In humans, regions along the bilateral STS, including posterior, mid and anterior STS (p/m/aSTS), extending to the TP, and also the right IFC were more sensitive to human than to either dog or non-vocal sounds. There was also a difference in the relation of the non-preferred dog and non-vocal sounds across the temporal subregions in humans. While in pSTS and mSTS the response to dog sounds was between that to human and non-vocal sounds, aSTS and TP regions showed no preference for dog compared to non-vocal sounds: repeated measures ANOVAs with factors anteriority (pSTS; mSTS; (right) aSTS; TP) and sound type (dog; non-vocal) showed a significant interaction of the two factors for each hemisphere (left:  $F(2,42)=45.386$ ,  $p<.001$ ; right:  $F(3,63)=42.491$ ,  $p<.001$ ). We also looked for regions responding stronger to heterospecific vocalizations than to other sound types. In dogs, no regions showed stronger responses to human than either dog or non-vocal sounds. In humans, only the subcortical MGB, but no temporal regions, showed greater sensitivity to dog than either human or non-vocal sounds.

These findings provide the first evidence for the existence of voice areas in dogs, or in any non-primate brain. The only bilateral conspecific-prefering region in dogs is near the TP, i.e. the ventral part of the caudal SG, extending to the SF. In humans, the anterior voice areas (i.e., aSTS, and TP) are special in that, similarly to dog voice areas, they respond most strongly to conspecific sounds, but do not respond stronger to heterospecific than to non-vocal sounds. These similarly located (i.e., near the TP) auditory cortex regions thus appear to be functionally analogous in dog and human brains. Anterior temporal and TP regions have been implied in conspecific vocalization processing in both non-human primates [7–10, 32, 33] and humans [6, 34, 35]. More specifically, these regions have been implied in voice identity processing, a key function of voice areas [6, 35, 36]. While claims about exact anatomical correspondences and therefore about homologies across dog and human brain regions are difficult to make and are beyond the scope of this paper, a plausible interpretation of our findings is that conspecific-preference in these auditory regions is an evolutionarily ancient function across mammalian orders, although convergent evolution [37] is an alternative. At the very least, our results show that, similarly to primates, conspecific vocalizations have a special status in the dog brain.

In humans, consistent with previous reports [5, 6, 35], the temporal voice areas involved not only anterior regions, but also the mid and posterior STS. Here we showed that pSTS and mSTS, unlike aSTS and TP, prefer heterospecific (dog) sounds to non-vocal sounds (cf. [38]). This suggests that pSTS and mSTS are not strictly conspecific-specific, but rather tuned to familiar, relevant vocal sounds in general, an interpretation possibly also supported by a report finding no preference for conspecific compared to human vocalizations in the macaque mSTS [32]. Replicating earlier findings, we also found conspecific-preference in the human right IFC [39, 40], another region implied in voice identity processing [40, 41]. Finally, we found that, in humans, the subcortical MGB, previously implied in processing rapidly varying

spectrotemporal features of human vocal sounds [42], responded stronger to dog sounds than other sound types.

We also tested, in a series of parametric modulation analyses, whether the emotional valence of vocalizations is reflected in brain responses, and how such responses are modulated by acoustical cues. Vocal stimuli were blocked by valence scores (ranging from highly negative to highly positive). The affective context valence of the dog vocalization recordings was found to covary with human emotional valence ratings of these sounds, suggesting that human ratings represent a fairly good evaluation of the animal's affective state ([16], and Supplemental Experimental Procedures). Block-averaged valence scores were then used as parametric modulators to test whether auditory brain activity covaries with emotional content. Specifically, we tested whether emotional vocalizations that are perceived as more positive (in a parametric manner) elicit greater (or smaller) neural responses (for details, see Supplemental Experimental Procedures).

Emotional valence-sensitive regions were identified in both dogs and humans (Figure 3). These regions all responded stronger to more positive vocalizations – we found no regions responding stronger to more negative vocalizations. In dogs, we found that an auditory region in the right cESG, close to the primary auditory cortex, was sensitive to emotional valence, for both dog and human vocalizations (Table S3). No emotional modulation effect was found in the corresponding cESG region of the left hemisphere, with a significant difference across hemispheres ( $T(10)=2.234$ ,  $p<.05$ , see Supplemental Experimental Procedures). We found a similar, but weaker modulatory effect of emotional valence in a bilateral rostral SG region for human but not dog vocalizations. In humans, an analogous effect was found: neural activity in the auditory cortex, with a maximum in the mSTS, increased with the perceived emotional positivity of vocalizations. This emotional modulation effect was present bilaterally for both dog and human vocal stimuli, with only a tendency for a right hemisphere bias ( $T(21)=1.879$ ,



$p=.074$ ). The human mSTS has been implied in the extraction of social or affective salient signals from conspecific vocalizations [11]. Our results show that the same mechanism may be used to extract affective information from heterospecific vocalizations. Furthermore, dogs appear to use a similar mechanism, localized in the cESG, for extracting vocal emotional information from either conspecifics or humans.

Additionally, we tested how auditory regions in each species are modulated by acoustic parameters relevant for emotional processing. In a related paper [16] we already established that perceived emotional valence and intensity of these vocal stimuli covary with a basic temporal cue (call length) and a basic spectral cue (fundamental frequency, F0) respectively. Specifically, emotional valence increases with decreasing call length, while emotional intensity increases with increasing F0. Here we found that auditory regions are parametrically modulated (1) by call length in dogs (activity decreased with increasing call length in the bilateral mESG and also in the subcortical MGB), and (2) by F0 in both species (activity decreased with increasing F0, in the right mESG and in a left rSG region in dogs, and in the m/pSTS in humans) (Table S3). In both dogs and humans, the right auditory cortex peaks for valence-sensitivity and each acoustic parametric effect were within a 16 mm distance, and close to primary auditory regions.

These findings suggest that acoustical cues related to vocal emotional valence are processed similarly in the dog and human auditory cortex. The involvement of a relatively early stage in the processing hierarchy in both species indicates that valence sensitivity at least partly reflects sensitivity in both species to acoustic parameters that convey emotions through voice. This is consistent with earlier human findings that imply emotional voice-sensitive regions in the mSTS and pSTS in processing both temporal [43] and spectral [12, 13, 31, 44] cues.

These discoveries suggest that the extraction of emotional information from voices is an important stage of the vocal emotion processing hierarchy, and is supported by functionally analogous auditory brain regions near the primary auditory cortex in dogs and humans. These results expand earlier findings that dogs react similarly to some emotional state changes of other dogs and humans [45], and that humans recruit similar brain regions to process human and animal affective vocalizations [14]. These results also demonstrate that right hemisphere dominance in vocal emotion processing, while debated in humans [46], is present in dogs, suggesting that behavioural lateralization effects in dog auditory processing [47] may be caused primarily by modulation of right hemisphere activity.

This fMRI study compared, for the first time, two phylogenetically distant mammalian species under almost identical experimental conditions. Our results suggest common functions in dog and human voice processing. We presented evidence that voice areas preferring conspecific vocalizations exist not only in primates, but also in dogs, and that, as in non-human primates [7–10, 32, 33] and humans [6, 34, 35], the dog voice areas involve bilateral TP regions. This evidence opens up the possibility that voice areas may have a longer evolutionary history than previously proposed ([10]), dating back to the common ancestor of dogs and humans some 100 million years ago [4], although convergent evolution cannot be excluded [37]. We also identified similarly located (i.e. near the primary auditory cortex) regions sensitive to emotional valence in vocalizations in both species, and showed that this valence sensitivity involves keeping track of basic acoustic cues that mediate vocal emotions. This may be the first direct evidence suggesting that voice processing in mammalian listeners corresponds to the structural-functional organisation of vocalizations [15] and forms the basis for using key acoustic features for cross-specific call recognition.

**Supplemental Information** consists of one figure, three tables, one video, one audio item and Supplemental Experimental Procedures.

### **Acknowledgements**

The study was supported by The Hungarian Academy of Sciences (F01/031), the Hungarian Scientific Research Fund (OTKA K100695), and by the ESF Research Networking Programme "CompCog": The Evolution of Social Cognition ([www.compcog.org](http://www.compcog.org)) (06-RNP-020). We thank Gábor Rudas and the MR Research Centre of the Semmelweis University Budapest; József Topál, Viktor Gál, Zita Polgár and four anonymous reviewers for comments on the manuscript; Ferenc Szalay for advice on dog brain anatomy; and Katalin Czigány, Borbála Ferenczy, Bernadett Miklósi, and all owners for training the dogs.

## References

1. Thalmann, O., Shapiro, B., Cui, P., Schuenemann, V. J., Sawyer, S. K., Greenfield, D. L., Germonpré, M. B., Sablin, M. V., López-Giráldez, F., Domingo-Roura, X., et al. (2013). Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science* *342*, 871–4. DOI: 10.1126/science.1243650.
2. Miklósi, A., and Topál, J. (2013). What does it take to become “best friends”? Evolutionary changes in canine social competence. *Trends Cogn. Sci.* *17*, 287–94. DOI: 10.1016/j.tics.2013.04.005.
3. Pongrácz, P., Molnár, C., and Miklósi, Á. (2010). Barking in family dogs: an ethological approach. *Vet. J.* *183*, 141–7. DOI: 10.1016/j.tvjl.2008.12.010.
4. Springer, M. S., Murphy, W. J., Eizirik, E., and O’Brien, S. J. (2003). Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 1056–61. DOI: 10.1073/pnas.0334222100.
5. Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* *403*, 309–12. DOI: 10.1038/35002078.
6. Andics, A., McQueen, J. M., Petersson, K. M., Gál, V., Rudas, G., and Vidnyánszky, Z. (2010). Neural mechanisms for voice recognition. *Neuroimage* *52*, 1528–40. DOI: 10.1016/j.neuroimage.2010.05.048.
7. Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., and Logothetis, N. K. (2008). A voice region in the monkey brain. *Nat. Neurosci.* *11*, 367–74. DOI: 10.1038/nn2043.

8. Perrodin, C., Kayser, C., Logothetis, N. K., and Petkov, C. I. (2011). Voice cells in the primate temporal lobe. *Curr. Biol.* *21*, 1408–15. DOI: 10.1016/j.cub.2011.07.028.
9. Poremba, A., Malloy, M., Saunders, R. C., Carson, R. E., Herscovitch, P., and Mishkin, M. (2004). Species-specific calls evoke asymmetric activity in the monkey's temporal poles. *Nature* *427*, 448–51. DOI: 10.1038/nature02268.
10. Gil-da-Costa, R., Lopes, M. A., Martin, A., Fritz, J. B., Braun, A. R., and Munoz, M. (2006). Species-specific calls activate homologs of Broca's and Wernicke's areas in the macaque. *Nat. Neurosci.* *9*, 1064–70. DOI: 10.1038/nn1741.
11. Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., and Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci.* *8*, 145–6. DOI: 10.1038/nn1392.
12. Ethofer, T., Van De Ville, D., Scherer, K. R., and Vuilleumier, P. (2009). Decoding of emotional information in voice-sensitive cortices. *Curr. Biol.* *19*, 1028–33. DOI: 10.1016/j.cub.2009.04.054.
13. Overath, T., Kumar, S., Stewart, L., von Kriegstein, K., Cusack, R., Rees, A., and Griffiths, T. D. (2010). Cortical mechanisms for the segregation and representation of acoustic textures. *J. Neurosci.* *30*, 2070–6. DOI: 10.1523/JNEUROSCI.5378-09.2010.
14. Belin, P., Fecteau, S., Charest, I., Nicastro, N., Hauser, M. D., and Armony, J. L. (2008). Human cerebral response to animal affective vocalizations. *Proc. R. Soc. B Biol. Sci.* *275*, 473–81. DOI: 10.1098/rspb.2007.1460.
15. Morton, E. S. (1977). On the occurrence and significance of motivation - structural rules in some bird and mammal sounds. *Am. Nat.* *111*, 855–869.

16. Faragó, T., Andics, A., Devecseri, V., Kis, A., Gácsi, M., and Miklósi, A. (2014). Humans rely on the same rules to assess emotional valence and intensity in conspecific and dog vocalizations. *Biol. Lett.* *10*, 20130926. DOI: 10.1098/rsbl.2013.0926.
17. August, P. V., and Anderson, J. G. T. (1987). Mammal Sounds and Motivation-Structural Rules: A Test of the Hypothesis. *J. Mammal.* *68*, 1–9.
18. Lemasson, A., Remeuf, K., Rossard, A., and Zimmermann, E. (2012). Cross-taxa similarities in affect-induced changes of vocal behavior and voice in arboreal monkeys. *PLoS One* *7*, e45106. DOI: 10.1371/journal.pone.0045106.
19. Fichtel, C., Hammerschmidt, K., and 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. DOI: 10.1163/15685390151067094.
20. 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. DOI: 10.1121/1.1568942.
21. Scheumann, M., Roser, A.-E., Konerding, W., Bleich, E., Hedrich, H.-J., and Zimmermann, E. (2012). Vocal correlates of sender-identity and arousal in the isolation calls of domestic kitten (*Felis silvestris catus*). *Front. Zool.* *9*, 36. DOI: 10.1186/1742-9994-9-36.
22. Graham, M. A., and Noonan, M. (2010). Call Types and Acoustic Features Associated with Aggressive Chase in the Killer Whale (*Orcinus orca*). *Aquat. Mamm.* *36*, 9–18. DOI: 10.1578/AM.36.1.2010.9.

23. Briefer, E. F. (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.
24. Taylor, A. M., and 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.
25. Owings, D. H., and Morton, E. S. (1998). *Animal vocal communication: A new approach* (Cambridge: Cambridge University Press) DOI: 10.1017/CBO9781139167901.
26. Tóth, L., Gácsi, M., Miklósi, Á., Bogner, P., and Repa, I. (2009). Awake dog brain magnetic resonance imaging. *J. Vet. Behav. Clin. Appl. Res.* 4, 50. DOI: 10.1016/j.jveb.2008.09.021.
27. Berns, G. S., Brooks, A. M., and Spivak, M. (2012). Functional MRI in Awake Unrestrained Dogs. *PLoS One* 7, e38027. DOI: 10.1371/journal.pone.0038027.
28. Heffner, H. (1978). Effect of auditory cortex ablation on localization and discrimination of brief sounds. *J. Neurophysiol.* 41, 963–76.
29. Stepień, I., Stepień, L., and Lubińska, E. (1990). Function of dog's auditory cortex in tests involving auditory location cues and directional instrumental response. *Acta Neurobiol. Exp. (Wars).* 50, 1–12.
30. Szwejkowska, G., and Sychowa, B. (1971). The effects of lesions of auditory cortex on discrimination of sound localization in dog. *Acta Neurobiol. Exp. (Wars).* 31, 237–50.

31. Von Kriegstein, K., Smith, D. R. R., Patterson, R. D., Ives, D. T., and Griffiths, T. D. (2007). Neural representation of auditory size in the human voice and in sounds from other resonant sources. *Curr. Biol.* *17*, 1123–8. DOI: 10.1016/j.cub.2007.05.061.
32. Joly, O., Pallier, C., Ramus, F., Pressnitzer, D., Vanduffel, W., and Orban, G. a (2012). Processing of vocalizations in humans and monkeys: A comparative fMRI study. *Neuroimage* *62*, 1376–1389. DOI: 10.1016/j.neuroimage.2012.05.070.
33. Joly, O., Ramus, F., Pressnitzer, D., Vanduffel, W., and Orban, G. a (2012). Interhemispheric differences in auditory processing revealed by fMRI in awake rhesus monkeys. *Cereb. Cortex* *22*, 838–53. DOI: 10.1093/cercor/bhr150.
34. Belin, P., and Grosbras, M.-H. (2010). Before speech: cerebral voice processing in infants. *Neuron* *65*, 733–5. DOI: 10.1016/j.neuron.2010.03.018.
35. Von Kriegstein, K., and Giraud, A.-L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage* *22*, 948–55. DOI: 10.1016/j.neuroimage.2004.02.020.
36. Von Kriegstein, K., Eger, E., Kleinschmidt, A., and Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cogn. Brain Res.* *17*, 48–55. DOI: 10.1016/S0926-6410(03)00079-X.
37. Fitch, W. T., Huber, L., and Bugnyar, T. (2010). Social cognition and the evolution of language: constructing cognitive phylogenies. *Neuron* *65*, 795–814. DOI: 10.1016/j.neuron.2010.03.011.



38. Fecteau, S., Armony, J. L., Joanette, Y., and Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *Neuroimage* 23, 840–8. DOI: 10.1016/j.neuroimage.2004.09.019.
39. Fecteau, S., Armony, J. L., Joanette, Y., and Belin, P. (2005). Sensitivity to voice in human prefrontal cortex. *J. Neurophysiol.* 94, 2251–4. DOI: 10.1152/jn.00329.2005.
40. Andics, A., McQueen, J. M., and Petersson, K. M. (2013). Mean-based neural coding of voices. *Neuroimage* 79, 351–60. DOI: 10.1016/j.neuroimage.2013.05.002.
41. Latinus, M., Crabbe, F., and Belin, P. (2011). Learning-induced changes in the cerebral processing of voice identity. *Cereb. Cortex* 21, 2820–8. DOI: 10.1093/cercor/bhr077.
42. Von Kriegstein, K., Patterson, R. D., and Griffiths, T. D. (2008). Task-dependent modulation of medial geniculate body is behaviorally relevant for speech recognition. *Curr. Biol.* 18, 1855–9. DOI: 10.1016/j.cub.2008.10.052.
43. Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., and Ethofer, T. (2008). Cerebral processing of emotional prosody--influence of acoustic parameters and arousal. *Neuroimage* 39, 885–93. DOI: 10.1016/j.neuroimage.2007.09.028.
44. Latinus, M., and Belin, P. (2011). Human voice perception. *Curr. Biol.* 21, R143–5. DOI: 10.1016/j.cub.2010.12.033.
45. Custance, D., and Mayer, J. (2012). Empathic-like responding by domestic dogs (*Canis familiaris*) to distress in humans: an exploratory study. *Anim. Cogn.* 15, 851–9. DOI: 10.1007/s10071-012-0510-1.

46. Schirmer, A., and Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* *10*, 24–30. DOI: 10.1016/j.tics.2005.11.009.
47. Siniscalchi, M., Quaranta, A., and Rogers, L. J. (2008). Hemispheric specialization in dogs for processing different acoustic stimuli. *PLoS One* *3*, e3349. DOI: 10.1371/journal.pone.0003349.

## Figure legends

**Figure 1.** Steps of positioning a dog in the fMRI scanner. a) Dog lying on scanner bed, being rewarded with food and socially by the owner. b) As part of the model-rival training procedure, another dog is observing as the tested dog is praised while receiving ear-phones from an experimenter. c) When the upper element of the coil is fixed with stripes on the top of the dog's head, the scanner bed is moved to the scanning position. See also Supplemental Experimental Procedures, Figure S1 and Video S1.

**Figure 2.** Auditory regions in dogs and humans. a) Schematic representations of sound-sensitive perisylvian regions in dogs and humans, superimposed on rendered brains. Dog abbreviations are c: caudal, m: middle, r: rostral, ESG: ectosylvian gyrus, ESS: ectosylvian sulcus, SF: Sylvian fissure, SG: Sylvian gyrus, SSS: suprasylvian gyrus, TP: temporal pole. Human abbreviations are a: anterior, m: mid, p: posterior, IFC: inferior frontal cortex, SF: Sylvian fissure, STS: superior temporal sulcus, TP: temporal pole. b) Auditory regions as determined by the all sounds vs silence contrast in dogs and humans, thresholded at  $p < .001$ , FWE-corrected at the cluster level, using the uncorrected voxel threshold  $p < .001$  for dogs (in a whole volume search space of  $90 \text{ cm}^3$ ), and  $p < .00001$  for humans (in a whole volume search space of  $1277 \text{ cm}^3$ ). Colour heat maps indicate t-values, superimposed on rendered brains and selected axial slices. c) The same auditory maps as in b, colour code refers to the sound type that elicited the maximal response in each voxel. See also Table S1 and Audio S1.

**Figure 3.** a) Species-preference and emotional valence sensitivity for vocalizations in dogs and humans, superimposed on rendered brains. Activity maps are thresholded at  $p < .005$  for dogs, and at  $p < .0005$  (in clusters of at least 10 voxels) for humans. Regions with human preference (a conjunction of human vocal  $>$  dog vocal and human vocal  $>$  non-vocal; red), dog preference (a conjunction of dog vocal  $>$  human vocal and dog vocal  $>$  non-vocal; blue),

human valence sensitivity (a positive parametric effect of valence of human vocal sounds; yellow) and dog valence sensitivity (a positive parametric effect of valence of dog vocal sounds; purple) are shown. b) Parameter estimates for voice area peaks. Bars represent beta weights for each sound type. Stripe bars: left hemisphere peaks, filled bars: right hemisphere peaks. Error bars indicate s.e.m. See also Table S2 and S3.

Figure 1  
[Click here to download high resolution image](#)



Figure 2  
[Click here to download high resolution image](#)

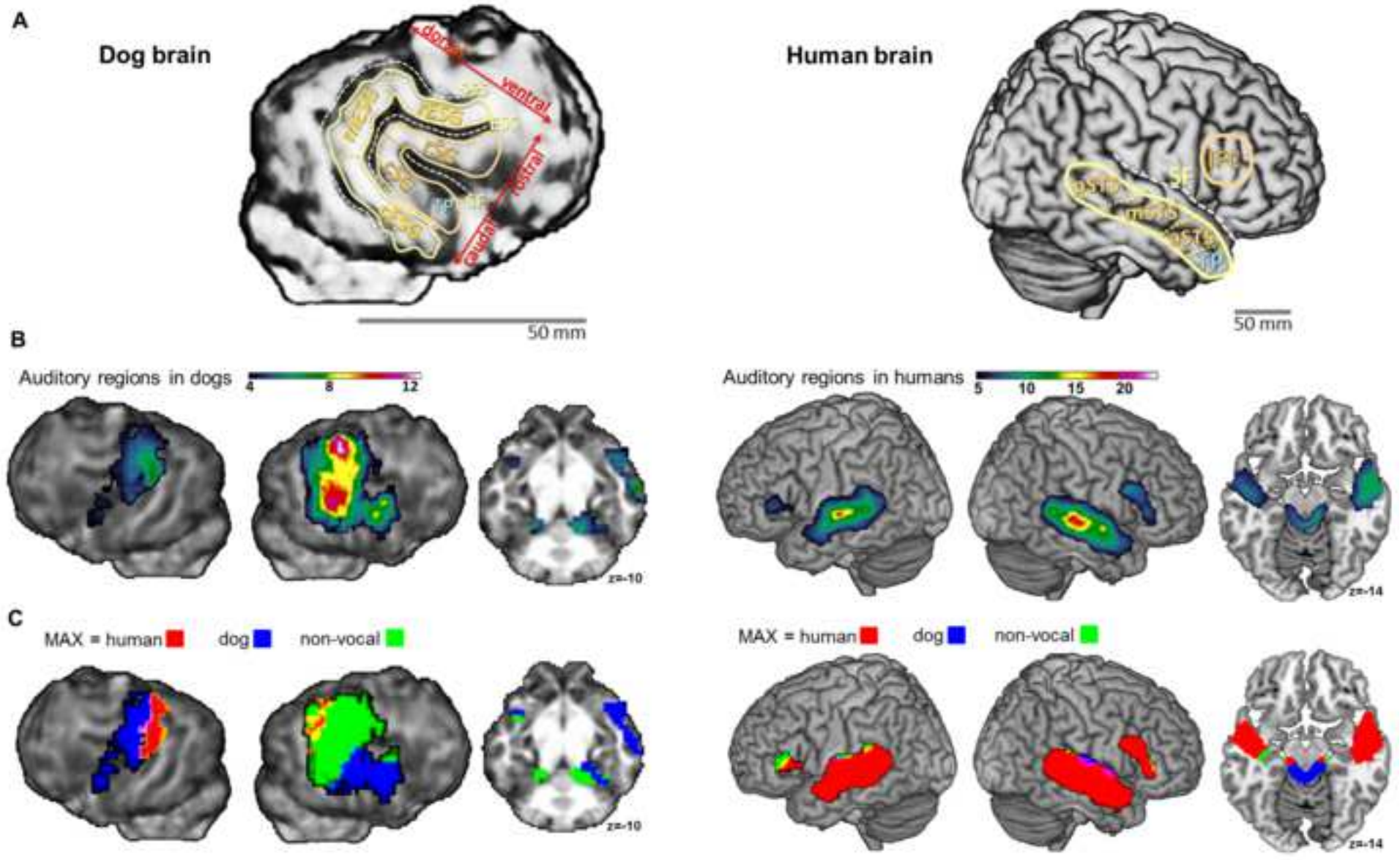
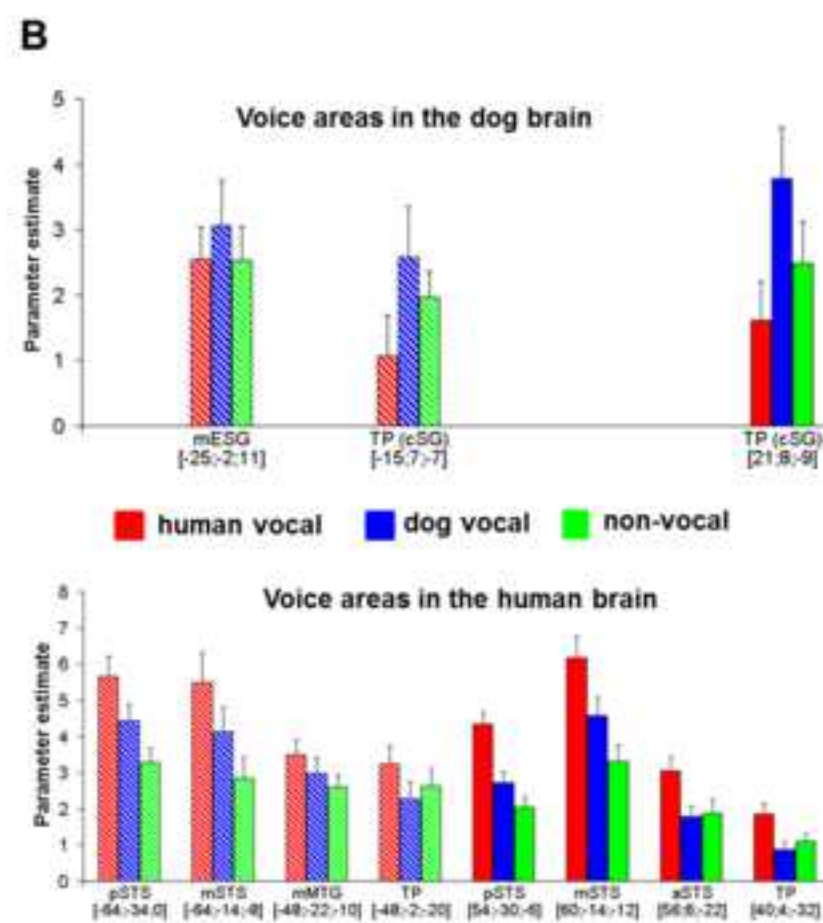
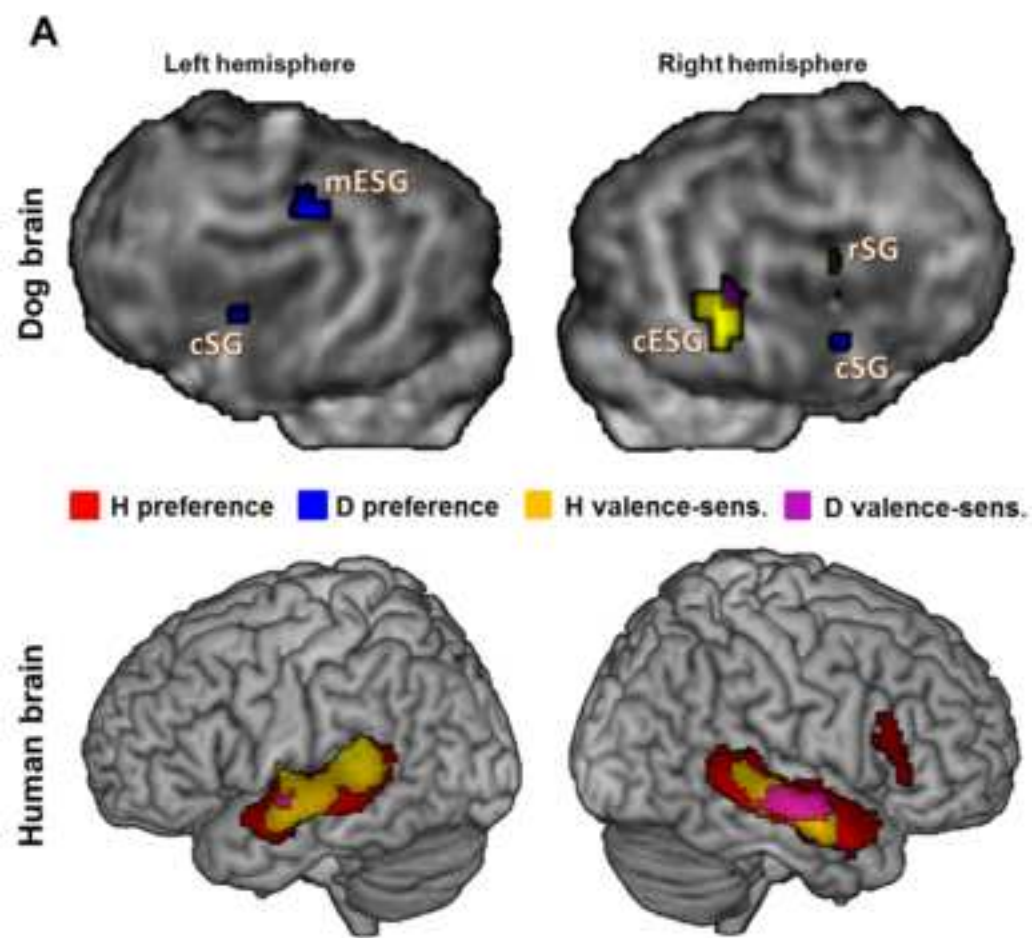


Figure 3  
[Click here to download high resolution image](#)



## **Inventory of Supplemental Information**

(Andics et al., Voice-sensitive regions in the dog and human brain are revealed by comparative fMRI, Ms. No. CURRENT-BIOLOGY-D-13-01192R1)

**Figure S1.** Individual dog movement parameters, related to Figure 1.

**Table S1.** Auditory brain regions in dogs and humans, related to Figure 2.

**Table S2.** Species-preference effects, related to Figure 3.

**Table S3.** Emotional valence sensitivity effects, related to Figure 3.

**Video S1.** Short video showing a trained dog going into the scanner, related to Figure 1.

[https://www.dropbox.com/s/v5o7hd706n5qd1m/Andics\\_Supplementary\\_Video.zip](https://www.dropbox.com/s/v5o7hd706n5qd1m/Andics_Supplementary_Video.zip)

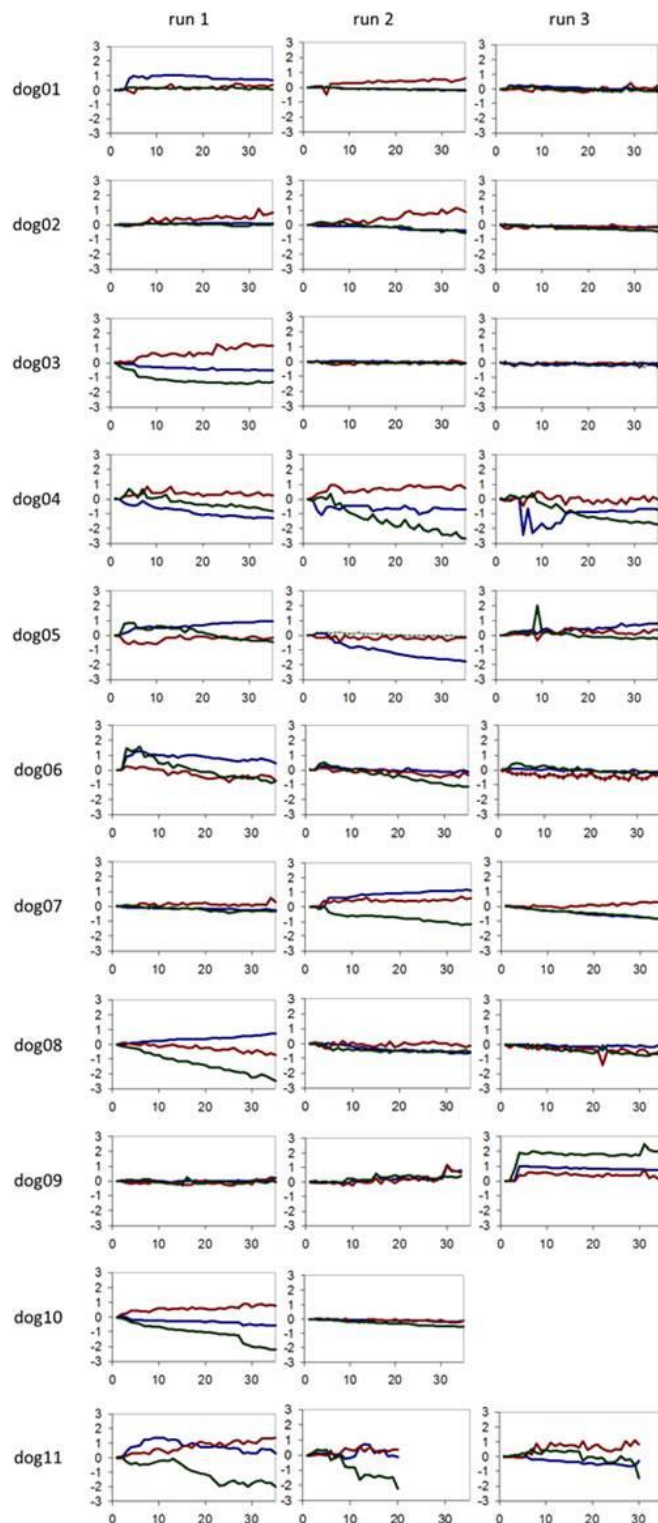
**Audio S1.** All human vocal (hum), dog vocal (dog), and non-vocal (noi) stimulus blocks – related to Figure 2.

[https://www.dropbox.com/s/gw3zo42v4wt50lh/Andics\\_Supplementary\\_Audio.zip](https://www.dropbox.com/s/gw3zo42v4wt50lh/Andics_Supplementary_Audio.zip)



## Supplemental Data

**Figure S1.** Individual dog movement parameters, related to Figure 1. Each row shows movements of one subject during the three runs along each of the three translation directions. (Rotations in any direction were below .1 degree and are not shown.) x axis: volume nr (each corresponds to 10 s in time). y axis: mm.



**Table S1.** Auditory brain regions in dogs and humans, related to Figure 2.

| Brain region | Dog listeners |     |     |       |                              |        | Human listeners |     |     |                              |       |        |
|--------------|---------------|-----|-----|-------|------------------------------|--------|-----------------|-----|-----|------------------------------|-------|--------|
|              | x             | y   | z   | T(10) | volume<br>(cm <sup>3</sup> ) | p(FWE) | x               | y   | z   | volume<br>(cm <sup>3</sup> ) | T(21) | p(FWE) |
| L aud cx     | -23           | -14 | 3   | 10.47 | 4.232                        | .000   | -44             | -26 | 2   | 40.92                        | 21.30 | .000   |
|              | -15           | 5   | -9  | 6.11  |                              |        | -58             | -40 | 6   |                              | 13.30 |        |
| R aud cx     | 25            | -10 | 3   | 12.33 | 6.176                        | .000   | 48              | -28 | 6   | 41.01                        | 23.07 | .000   |
|              | 23            | 8   | -3  | 8.53  |                              |        | 66              | -24 | 2   |                              | 15.32 |        |
| L IFC        |               |     |     |       |                              |        | -38             | 28  | 10  | 2.82                         | 8.76  | .000   |
|              |               |     |     |       |                              |        | -50             | 16  | 4   |                              | 5.55  |        |
| R IFC        |               |     |     |       |                              |        | 48              | 16  | 24  | 5.12                         | 9.08  | .000   |
|              |               |     |     |       |                              |        | 50              | 28  | 6   |                              | 7.75  |        |
| MGB          | 7             | -19 | -8  | 8.41  | 1.120                        | .000   | 10              | -32 | -14 | 4.92                         | 9.64  | .000   |
|              | -9            | -17 | -10 | 5.94  |                              |        |                 |     |     |                              |       |        |

Auditory clusters defined in whole-brain random effects analyses of the contrast all sounds [dog + human + non-vocal] > silence. Threshold for reporting was  $p < .001$  (FWE-corrected at the cluster level), using the voxel threshold  $p < .001$  for dogs and  $p < .00001$  (in clusters of at least 10 voxels) for humans. At most two peaks, at least 16 mm apart are reported.

**Table S2.** Species-preference effects, related to Figure 3.

| <b>Dog listeners</b>   | ROI                            | Brain region | x   | y   | z   | T(10) | p    |
|------------------------|--------------------------------|--------------|-----|-----|-----|-------|------|
| Human > dog, non-vocal | <i>no significant clusters</i> |              |     |     |     |       |      |
| Dog > human, non-vocal | R aud cx                       | caudal SG    | 21  | 8   | -9  | 3.7   | .002 |
|                        | L aud cx                       | caudal SG    | -15 | 7   | -7  | 3.41  | .003 |
|                        |                                | middle ESG   | -25 | -2  | 11  | 3.47  | .003 |
| <b>Human listeners</b> | ROI                            | Brain region | x   | y   | z   | T(21) | p    |
| Human > dog, non-vocal | R aud cx                       | post STS     | 54  | -30 | -6  | 10.47 | .000 |
|                        |                                | mid STS      | 60  | -14 | -12 | 8.65  | .000 |
|                        |                                | ant STS      | 56  | 6   | -22 | 7.35  | .000 |
|                        |                                | TP           | 40  | 4   | -32 | 7.07  | .000 |
|                        |                                | post STS     | 66  | -36 | 4   | 6.34  | .000 |
|                        |                                | post STS     | 40  | -26 | 4   | 4.26  | .000 |
|                        | L aud cx                       | mid STS      | -64 | -14 | -8  | 8.26  | .000 |
|                        |                                | post STS     | -64 | -34 | 0   | 6.47  | .000 |
|                        |                                | mid MTG      | -48 | -22 | -10 | 5.58  | .000 |
|                        |                                | post STS     | -54 | -48 | 4   | 5.43  | .000 |
|                        |                                | TP           | -48 | -2  | -20 | 4.84  | .000 |
|                        | R IFC                          |              | 46  | 22  | 22  | 4.66  | .000 |
|                        |                                |              | 56  | 26  | 4   | 4.30  | .000 |
| Dog > human, non-vocal | MGB                            |              | -14 | -34 | -16 | 4.41  | .000 |

‘Voice areas’ in dogs and humans, as determined by random effects of the first-level conjunctions human > dog and human > non-vocal, or dog > human and dog > non-vocal. Threshold for reporting was  $p < .005$  for dogs, and  $p < .0005$  (in clusters of at least 10 voxels) for humans. Brain search space was narrowed down to the auditory regions (see Figure 2 and Table S1). All peaks at least 16 mm apart are reported.

**Table S3.** Emotional valence sensitivity and acoustic effects, related to Figure 3.

| <b>Dog listeners</b>   |          | ROI                            | Brain region | x   | y   | z   | T(10) | p    |
|------------------------|----------|--------------------------------|--------------|-----|-----|-----|-------|------|
| Valence (dog)          | R aud cx |                                | caudal ESG   | 30  | -7  | -1  | 4.92  | .000 |
| Valence (human)        | R aud cx |                                | caudal ESG   | 28  | -9  | -7  | 4.56  | .001 |
|                        |          |                                | rostral SG   | 19  | 6   | 3   | 3.69  | .002 |
| Call length            | L aud cx |                                | rostral SG   | -16 | 5   | 3   | 3.33  | .004 |
|                        | R aud cx |                                | middle ESG   | 22  | -2  | 11  | 4.55  | .001 |
|                        | L aud cx |                                | middle ESG   | -21 | -4  | 7   | 5.51  | .000 |
| F0                     | MGB      |                                |              | -1  | -14 | -2  | 7.81  | .000 |
|                        | R aud cx |                                | middle ESG   | 28  | 2   | 9   | 4.82  | .000 |
|                        | L aud cx |                                | rostral SG   | -17 | 8   | -1  | 4.14  | .001 |
| <b>Human listeners</b> |          | ROI                            | Brain region | x   | y   | z   | T(21) | p    |
| Valence (dog)          | R aud cx |                                | mid STS      | 66  | -16 | -8  | 5.78  | .000 |
|                        | L aud cx |                                | mid STS      | -54 | -8  | 6   | 4.07  | .000 |
| Valence (human)        | R aud cx |                                | mid STS      | 64  | -12 | -12 | 7.29  | .000 |
|                        |          |                                | post STS     | 62  | -26 | -2  | 5.73  | .000 |
|                        |          |                                | MTG          | 44  | -34 | -2  | 5.09  | .000 |
|                        | L aud cx |                                | post STS     | -58 | -40 | 8   | 5.97  | .000 |
| Call length            |          | <i>no significant clusters</i> |              |     |     |     |       |      |
| F0                     | R aud cx |                                | mid STS      | 62  | -12 | -14 | 6.39  | .000 |
|                        |          |                                | post STS     | 44  | -38 | -2  | 4.73  | .000 |
|                        |          |                                | mid STS      | -64 | -10 | -10 | 5.76  | .000 |

Parametric effects of vocal sounds' emotional valence, call length and F0 are shown. For valence, only positive covariations were found (i.e., stronger activity for more positive vocalizations). For acoustic parameters, only negative covariations were found (i.e., stronger activity for smaller call length or F0 scores). Threshold for reporting was  $p < .005$  for dogs, and  $p < .0005$  (in clusters of at least 10 voxels) for humans. Brain search space was narrowed down to the auditory regions as determined in Table S1. All peaks at least 16 mm apart are reported.

**Video S1.** Short video showing a trained dog going into the scanner, related to Figure 1.

[https://www.dropbox.com/s/v5o7hd706n5qd1m/Andics\\_Supplementary\\_Video.zip](https://www.dropbox.com/s/v5o7hd706n5qd1m/Andics_Supplementary_Video.zip)

**Audio S1.** All human vocal (hum), dog vocal (dog), and non-vocal (noi) stimulus blocks – related to Figure 2. Block number within stimulus type corresponds to perceived emotional positivity (01: most negative, 24: most positive).

[https://www.dropbox.com/s/gw3zo42v4wt50lh/Andics\\_Supplementary\\_Audio.zip](https://www.dropbox.com/s/gw3zo42v4wt50lh/Andics_Supplementary_Audio.zip)

## **Supplemental Experimental Procedures**

### **Participants**

There were 11 dog participants (6 border collies and 5 golden retrievers; age  $3.4 \pm 2.0$  years; 6 male, 5 female) and 22 human participants (11 male, 11 female; age  $22.5 \pm 2.8$ ). Human participants and the owners of dog participants volunteered to take part in the project without monetary compensation and gave written informed consent. Experimental procedures met the national and European guidelines for animal care and were approved by the local ethical committee (Pest Megyei Kormányhivatal Élelmiszerlánc-Biztonsági és Állategészségügyi Igazgatósága XIV-I-001/520-4/2012, Budapest, Hungary).

### **Dog training**

All dogs lived with their owners as family pets and had various types of pre-training independent of the research (nothing, basic, agility, obedience). In order to prepare the dogs for the awake fMRI testing, a special step-by-step training procedure was planned by MG (maintaining the required position on a table, getting used to the ear-phones and strips/coil, being habituated to the noise and vibration of the scanner). The training was based on two methods: conditioning and social learning. First, positive reinforcement techniques were applied, that is, dogs were continuously food-rewarded, praised, and stroked for the desired behavior. In addition, especially in the beginning of the training at the scanner, some aspects of the “Model/Rival” training method [S1] were adopted. Novice dogs were allowed to participate off-leash in the scanner room during the training session of a familiar dog (another individual from the same household or a “friend” from the dog school). When the model (the dog in the scanner) was praised and rewarded by the trainer and both owners, the novice dog was ignored. During the training sessions, we did not apply any restraints to keep the dogs in the desired position, and they could leave the tube any time. Before their first visit at the scanner, dogs needed on average 12 sessions with a trainer (range: 5-20). On average 7 sessions were needed (range: 5-9) with the scanner before the first functional measurements.

## Stimuli

We selected 96 human nonlinguistic vocalizations, 96 dog vocalizations, and 96 non-vocal environmental sounds. Human and non-vocal stimuli were collected from available databases [S2–S4], dog stimuli were selected from an in-house collected database. With both dog and human stimuli, our aim was to cover the natural repertoire as widely as possible.

Dog vocalizations were recorded from over 50 dogs in various social contexts (during play, petting, resting, asking for toy, begging for food, requesting activity, excitement before walking, food or territory guarding, being threatened, in separation, schutzhund training), and contained several call types [S5] (yelp, bark, growl, grunt, moan, whine and pant). These dog vocalizations were also scored based on the valence of their recording contexts. Situations which typically evoke discomfort, avoidance or stress in dogs were assigned the score -1 (threatening, separation, hunger, guarding and schutzhund), while contexts evoking comfort or attraction were assigned +1 (play, petting, greeting, excitement). The ambiguous or neutral contexts (resting, activity requesting) were marked with 0.

Human stimuli were from various male, female, infant and child vocalizers. These nonverbal vocalizations were either spontaneous or acted to express basic emotions (happiness, fear, anger, disgust, sadness, surprise, pleasure, and pain), and contained several call categories (laugh, shout, cry, scream, erotic moan, moan, cough, retch, sigh, yawn, and nonsense babbling).

All vocal sounds were previously scored for perceived emotional valence and emotional intensity by 36 independent human listeners (see [S6]). In case of dog vocalizations, the valence ratings were validated with the context valence categorization. The comparison of the averaged valence ratings between positive (+1) and negative (-1) contexts showed that sounds originating from positive contexts were indeed rated significantly more positive (Mann-Whitney test:  $U=789$ ;  $p=0.001$ ). Stimuli were digitized at a 16 bit/22 kHz sampling rate and were equalized for -26dB RMS using Adobe Audition CS5.5.

## Design and procedure

Identical stimulus designs were applied for both dog and human participants. Stimuli were grouped into 4-stimulus-long blocks. This blocking was random for the non-vocal sounds, and was based on the valence scores for the vocal sounds (as in [S7]). For both human and dog vocalizations, block01 contained those 4 stimuli that had the 4 most negative scores, block02 contained the next 4 negative stimuli, and so on, with block24 containing the 4 most positively scored stimuli. Thus, 24 blocks were created for each stimulus type. Stimulus blocks were matched for number of sources, duration, and overall energy. All stimuli were at most 2 s long. Stimulus onsets within the block were at 0, 2, 4 and 6 s. All blocks were at most 8 s long (see Audio S1).

During the fMRI tests, four conditions were used, each corresponding to a stimulus block type. The block types were: human vocalizations, dog vocalizations, non-vocal sounds, and silence. Silence blocks were also added as a baseline condition. The fMRI test was split into three runs, each containing 8 blocks per type, distributed evenly with respect to the average valence scores of the vocal blocks, and randomly for the non-vocal blocks. Consecutive blocks were always from a different type. Stimulus order within blocks, block order within runs, and run order were all pseudorandomized, and varied across participants.

Every stimulus block (8 s) was presented in silence and was followed by a 2 s long volume acquisition which was then immediately followed by the next block (see [S4] for a similar protocol). Stimulus block onset asynchrony was therefore 10 s. Each run contained 32 blocks. Two dummy scans were added to the beginning of each run, and 3 extra scans were added to the end, so the total length of a run was approximately 6 mins. Stimuli were presented at a 68 dB volume level.

Stimuli were controlled using Matlab (version 9.1) Psychophysics Toolbox 3 [S8]. During imaging, stimulus presentation was synchronized by a TTL trigger pulse with the data acquisition. Stimuli were delivered binaurally through MRI-compatible sound-attenuating headphones (MR Confon, Magdeburg, Germany) that were suitable to cover the ears of both dog and human participants.



## Data acquisition

MRI measurements were taken at the MR Research Centre of the Semmelweis University Budapest on a Philips Achieva 3 T whole body MR unit (Philips Medical Systems, Best, The Netherlands).

For dog participants, a Philips SENSE Flex Medium coil was used. This coil consisted of two elliptical elements that were 14 cm x 17 cm. One element was placed below the dog's head under the carpet on which the dog laid; the other element was fixed with plastic strips above the dog's head (see Figure 1 and Video S1). EPI-BOLD fMRI time series were obtained from 29 transverse slices covering the whole brain with a spatial resolution of 3.5×3.5×3.5 mm, including a 0.5 mm slice gap, using a single-shot gradient-echo planar sequence (ascending slice order; acquisition matrix 64×64; TR=10000 ms, including 2000 ms acquisition and 8000 ms silent gap; TE=36 ms; flip angle=90°). Each of the three runs included 35 volumes. In addition to the functional time series, a standard T1-weighted three-dimensional scan using a turbo-field echo (TFE) sequence with 180 slices covering the whole brain was collected for anatomical reference, with 1×1×1 mm spatial resolution.

For human participants, an eight-channel Philips SENSE head coil was used. EPI-BOLD fMRI time series were obtained from 29 transverse slices covering temporal lobes and the inferior part of the frontal lobes. All other parameters were set identical to those applied for the dog participants.

Dog participants were trained to remain still for the duration of the tests, were not restrained in any way, and could leave the setting by withdrawing their head at any time. Human participants were instructed to passively listen to the stimuli.

Dog participants were tested one run per session, with no more than two sessions per day. Runs with excessive motion and runs that the dog interrupted by withdrawing its head from the coil were not used. We continued test sessions with each dog until we obtained one structural run and three functional runs. The average number of test sessions (aiming at functional measurements) needed until the final set of runs was completed was 3.4 (range: 3-4). Human participants were tested in a single session.

## Data analysis

Image preprocessing and statistical analysis were performed using SPM8 ([www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm)). The functional EPI-BOLD images were realigned. For dogs, if head translation exceeded 3 mm or rotation exceeded 1 degree, the affected and all following volumes of the scan were excluded (30 of 33 functional runs were not affected by these exclusion criteria; one run of Dog10, volume 21-35 of a run of Dog11, and volume 35 of a run of Dog09 was excluded). The average of maximal movements per dog was below 1.74 mm for each translation direction, and below .018 degree for each rotation direction – realignment parameters for each dog participant are reported in Figure S1. The structural MR image was co-registered to the mean functional image and spatially normalized and transformed via SPM's standard nonlinear warping function with 16 iterations into a common anatomical space (using SPM Montreal Neurological Institute (MNI) T1 template for humans, and a selected golden retriever participant's smoothed (FWHM 2 mm) anatomical image for dogs). The origin of the dog brain space was positioned in the brain midline plane, at the most superior peak of the thalamus; negative to positive x, y, and z coordinates are in mm and, similarly to the MNI space for humans, denote left to right, posterior to anterior, and inferior to superior directions respectively. The resulting transformation matrix was then applied to all functional images. Normalized functionals were then spatially filtered by convolving the images with an isotropic 3-D Gaussian kernel. FWHM was 6 mm for dogs and 10 mm for humans – kernel size difference corresponds approximately to the brain size differences.

The fMRI data were analyzed using a general linear model and statistical parametric mapping [S9]. We constructed condition regressors for each run and for each block type: human vocalizations, dog vocalizations, non-vocal sounds, and silence. Conditions were modeled as 8 s long blocks. For the vocalization conditions, average perceived emotional valence and its absolute value per trial were added to the model as parametric modulators. Realignment regressors for each run were also included to model potential movement artifacts. Average total movement of dog participants, as calculated from the three translational directions, did not differ among the four (human/dog/non-vocal/silence) conditions (ANOVA,  $F(3,40)=.141$ ,  $p=.935$ ) (also see Figure S1). A high-pass filter with a cycle-cutoff of 128 s was implemented in the design to remove low-frequency signals. Regressors were convolved with the canonical haemodynamic response function of SPM. For both dogs and humans, single-subject fixed effect analyses were followed by whole-volume random effects analyses on the group level.

Auditory regions were defined functionally, using the all sounds vs silence contrast (thresholded at  $p < .001$ , FWE-corrected at the cluster level, using the uncorrected voxel threshold  $p < .001$  for dogs and  $p < .00001$  for humans, see Figure 2). The involvement of the MGB in the activated subcortical network was confirmed by visual inspection at various contrast thresholds. Note that the peak coordinates for MGB in humans were almost identical to those in [S10]. For higher-level contrasts, brain search space was narrowed down to the auditory regions (see Table S1). Threshold for reporting for all higher-level contrasts was  $p < .005$  for dogs, and  $p < .0005$  (in clusters of at least 10 voxels) for humans. All peaks at least 16 mm apart are reported. Effects of species preference in dogs and humans were determined by random effects analyses of the first-level conjunctions human > dog AND human > non-vocal, or dog > human AND dog > non-vocal, using the default minimum t statistic approach of SPM8 (Figure 3 and Table S2). Parametric effects of perceived emotional valence of vocal stimuli were tested in random effects parametric modulation analyses (Figure 3 and Table S3). Acoustic parameters' modulatory effects were estimated using separate general linear models, including condition regressors for each run and for each block type, and a parametric modulator for call length mean or F0 mean. (Call length here refers to the average length of the individual units within a sound, e.g. average bark length within a barking sequence, or average laugh length within a laughing sequence – see [S6] for details.).

The right-hemisphere auditory region showing maximal sensitivity to emotional valence for both human and dog sounds (i.e., 8-mm radius spheres around the peak [30; -7; -1] for dogs, and the peak [64; -12; -12] for humans) and their left-hemisphere counterparts were used in the hemispheric lateralization test, by comparing average parameter estimates in paired t-tests.

### **Supplemental References**

- S1. Pepperberg, I. M. (1981). Functional Vocalizations by an African Grey Parrot (*Psittacus erithacus*). *Z. Tierpsychol.* 55, 139–160. DOI: 10.1111/j.1439-0310.1981.tb01265.x.
- S2. Belin, P., Fillion-Bilodeau, S., and Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behav. Res. Methods* 40, 531–539. DOI: 10.3758/BRM.40.2.531.

- S3. Bradley, M. M., and Lang, P. J. (2007). *The International Affective Digitized Sounds (2nd Edition IADS-2): Affective ratings of sounds and instruction manual* (Gainesville, FL).
- S4. Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* *403*, 309–12. DOI: 10.1038/35002078.
- S5. Cohen, J. A., and Fox, M. W. (1976). Vocalizations in wild canids and possible effects of domestication. *Behav. Processes* *1*, 77–92. DOI: 10.1016/0376-6357(76)90008-5.
- S6. Faragó, T., Andics, A., Devecseri, V., Kis, A., Gácsi, M., and Miklósi, A. (2014). Humans rely on the same rules to assess emotional valence and intensity in conspecific and dog vocalizations. *Biol. Lett.* *10*, 20130926. DOI: 10.1098/rsbl.2013.0926.
- S7. Joly, O., Pallier, C., Ramus, F., Pressnitzer, D., Vanduffel, W., and Orban, G. a (2012). Processing of vocalizations in humans and monkeys: A comparative fMRI study. *Neuroimage* *62*, 1376–1389. DOI: 10.1016/j.neuroimage.2012.05.070.
- S8. Kleiner, M., Brainard, D., Pelli, D., and Ingling, A. (2007). What's new in Psychtoolbox-3. *Perception* *36*.
- S9. Friston, K. J., Ashburner, J. T., Kiebel, S., Nichols, T., and Penny, W. eds. (2007). *Statistical parametric mapping: the analysis of functional brain images* (London, UK: Academic Press Inc).
- S10. Von Kriegstein, K., Patterson, R. D., and Griffiths, T. D. (2008). Task-dependent modulation of medial geniculate body is behaviorally relevant for speech recognition. *Curr. Biol.* *18*, 1855–9. DOI: 10.1016/j.cub.2008.10.052.

Video S1

[Click here to download Supplemental Movie and Spreadsheet: REV\\_Andics\\_VideoS1.avi](#)