

HUMANS RECOGNIZE VOCAL EXPRESSIONS OF EMOTIONAL STATES UNIVERSALLY ACROSS SPECIES

PIERA FILIPPI^{1,2}, JENNA V. CONGDON³, JOHN HOANG³, DANIEL LIU
BOWLING⁴, STEPHAN REBER⁴, ANDRIUS PAŠUKONIS⁴, MARISA
HOESCHELE⁴, SEBASTIAN OCKLENBURG⁵, BART DE BOER¹, CHRISTOPHER
B. STURDY³, ALBERT NEWEN^{2,6}, ONUR GÜNTÜRKÜN^{2,5}

¹ Artificial Intelligence Laboratory, Vrije Universiteit Brussel, Belgium; ²Center for Mind, Brain and Cognitive Evolution Email: pie.filippi@gmail.com; ³Department of Psychology, University of Alberta, Canada; ⁴Department of Cognitive Biology, University of Vienna; ⁵Department of Biopsychology, Ruhr University, Bochum, Germany; ⁶Department of Philosophy II, Ruhr University, Bochum, Germany;

The perception of danger in the environment can induce physiological responses (such as a heightened state of arousal) in animals, which may cause measurable changes in the prosodic modulation of the voice (Briefer, 2012). The ability to interpret the prosodic features of animal calls as an indicator of emotional arousal may have provided the first hominins with an adaptive advantage, enabling, for instance, the recognition of a threat in the surroundings. This ability might have paved the way for the ability to process meaningful prosodic modulations in the emerging linguistic utterances.

Research has shown that humans are able to recognize different levels of arousal in mammalian calls. However, to our knowledge, no study has ever examined humans' cross-cultural ability to identify different arousal levels in calls of species belonging to several phylogenetically distant taxa, including, for instance, mammals and birds. We addressed this issue by developing a task in which human participants of three different cultures (Canadian, German, Mandarin) listened to ten pairs of vocalizations for each of nine different vertebrate taxa. We used amplitude-controlled calls from the following species: hourglass treefrog, American alligator, black-capped chickadee, common raven, domestic pig, giant panda, African elephant, Barbary macaque, and human.

Calls within each pair differed in arousal level, which was assessed based on the behavioral context of call production (Avey et al., 2011; Bowling et al., 2012; Fischer et al. 1995; Linhart et al., 2015; Reichert 2013; Stoeger et al., 2011, 2012). For each pair of vocalizations, participants were asked to identify the call with the higher level of arousal. Accuracy rate in identifying arousal in each species was higher than expected by chance in all three cultures. No significant differences were observed between cultures. This finding provides empirical support for Darwin's hypothesis on the universality of vocal emotional communication. In order to better understand the mechanisms underlying emotional intensity recognition in our set of calls, we investigated which acoustic parameters correlate with participants' correct responses. We performed this analysis in two steps. First, we identified two acoustic features measurable in all calls of our stimuli set: duration and a frequency related measure, the spectral center of gravity. Second, we calculated the duration and the spectral center of gravity ratio for each pair of calls and correlated these two feature comparisons with the percentage of correct responses across pairs. Our data indicate that the spectral center of gravity is the only feature found across our animal species calls, which significantly correlates with the ability to discriminate high arousal calls. Further work within this research paradigm will provide quantitative data on shared mechanisms involved in emotional vocalizations' production and perception across animal taxa, investigating the perception of arousal in nonhuman species. This may improve our understanding of the semantic value of prosody in animal communication, and of its role in the emergence of human language.

References

- Avey, M. T., Hoeschele, M., Moscicki, M. K., Bloomfield, L. L., & Sturdy, C. B. (2011). Neural correlates of threat perception: neural equivalence of conspecific and heterospecific mobbing calls is learned. *PloS one*, *6*, e23844.
- Bowling, D. L., Sundararajan, J., Han, S. E., & Purves, D. (2012). Expression of emotion in Eastern and Western music mirrors vocalization. *PLoS One*, *7*, e3194.
- Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. *Journal of Zoology*, *288*, 1-20.
- Fischer, J., Hammerschmidt, K., & Todt, D. (1995). Factors affecting acoustic variation in Barbary-macaque (*Macaca sylvanus*) Disturbance Calls. *Ethology*, *101*, 51-66.
- Linhart, P., Ratcliffe, V. F., Reby, D., & Špinka, M. (2015). Expression of Emotional Arousal in Two Different Piglet Call Types. *PloS one*, *10*(8), e0135414.

- Reichert, M. S. (2013). Sources of variability in advertisement and aggressive calling in competitive interactions in the grey treefrog, *Hyla versicolor*. *Bioacoustics*, *22*, 195-214.
- Stoeger, A. S., Charlton, B. D., Kratochvil, H., & Fitch, W. T. (2011). Vocal cues indicate level of arousal in infant African elephant roars. *The Journal of the Acoustical Society of America*, *130*(3), 1700-1710.
- Stoeger, A. S., Baotic, A., Li, D., & Charlton, B. D. (2012). Acoustic features indicate arousal in infant giant panda vocalisations. *Ethology*, *118*(9), 896-905.