

Journal Club

Editor's Note: These short, critical reviews of recent papers in the *Journal*, written exclusively by graduate students or postdoctoral fellows, are intended to summarize the important findings of the paper and provide additional insight and commentary. For more information on the format and purpose of the Journal Club, please see http://www.jneurosci.org/misc/ifa_features.shtml.

Neurocognitive Mechanisms for Vocal Emotions: Sounds, Meaning, Action

Nadine Lavan¹ and César F. Lima^{2,3}

¹Department of Psychology, Royal Holloway University of London, Egham, Surrey, TW20 0EX, United Kingdom, ²Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom, and ³Center for Psychology, University of Porto, 4200-135 Porto, Portugal
Review of Bestelmeyer et al.

Humans express emotions in many different ways. Facial expressions, body postures, or vocal cues, for instance, communicate a wealth of information about others' emotional states. While the adaptive significance of these multiple cues has long been acknowledged (Darwin, 1872/2009), facial expressions have historically received more research attention than expressions via other channels. The interest in vocal emotions is increasing, but mostly focused on speech prosody, i.e., voice modulations in speech. Vocal communication does, however, additionally encompass diverse nonverbal vocalizations, such as laughter, sobs, or screams. Accounting for these signals is crucial for a complete understanding of vocal emotions.

In a recent study published in *The Journal of Neuroscience*, Bestelmeyer et al. (2014) provide new knowledge on this issue by demonstrating that nonspeech vocalizations are processed in multiple steps involving distinct brain networks: bilateral auditory cortices responded to vocal-

izations' low-level acoustic features, and a wider network including the anterior insulae and prefrontal systems processed higher-order evaluative aspects. While current models of vocal affect perception predict such a multistep processing of vocal emotions (Schirmer and Kotz, 2006; Brück et al., 2011), empirical evidence for this has been relatively sparse and limited to speech prosody. These new findings thus form an important contribution to the field.

Multiple steps in the processing of vocal emotions

The findings of Bestelmeyer et al. (2014) are based on a forced-choice emotion categorization task of vocalizations that were morphed on continua between anger and fear. The task was performed by 19 participants in a magnetic resonance imaging scanner, using a continuous carry-over design (Aguirre, 2007) to examine how neural responses to a given vocalization were modulated by the features of the previously presented one: attenuated responses were expected when the preceding vocalization shared a dimension of interest (acoustic features, perceptual/evaluative aspects) versus when it differed (adaptation or "carry-over" effects). The experiment included six different sequences of 65 stimuli, each including eight items consisting of seven morph steps of a continuum between anger and fear, and a silent null event (the first item was repeated nine times per sequence, and the remain-

ing ones were repeated eight times; a given item was preceded and followed by every other items an equal number of times). The authors build upon previous behavioral evidence of auditory aftereffects for vocal expressions, which are suggestive of neural adaptation phenomena for vocal emotion categories (Bestelmeyer et al., 2010).

To separate low-level acoustic from higher-order evaluative processes, an index of "physical difference" was computed based on the absolute difference in morph steps between each vocalization and the preceding one. An index of "perceptual difference" was also computed, corresponding to the absolute difference in participants' subjective judgments of consecutive stimuli (i.e., the difference in the proportion of fear categorizations between each vocalization and the preceding one). In the analysis of the neuroimaging data, this perceptual index was orthogonalized to the physical one, allowing the investigators to examine adaptation effects produced by higher-order processes after accounting for the acoustic-related variance in neural responses. Isolating higher-order processes is an impressive feat of this study. However, as the authors acknowledge, completely isolating low-level processes was not possible: physical and perceptual differences are not independent. Thus, although consistent with the predictions of theoretical models (Schirmer and Kotz, 2006; Brück et al., 2011), the reported associations between

Received July 21, 2014; revised Aug. 13, 2014; accepted Aug. 13, 2014.

C.F.L. is supported by a postdoctoral fellowship from the Portuguese Foundation for Science and Technology (SFRH/BPD/77189/2011). We thank Dr. Carolyn McGettigan and Prof. Sophie Scott for their comments on an earlier draft of this article.

Correspondence should be addressed to either of the following: Nadine Lavan, Department of Psychology, Royal Holloway University of London, Egham, Surrey, TW20 0EX, United Kingdom, E-mail: nadine.lavan.2013@rhul.ac.uk; or César Lima, Center for Psychology, University of Porto, Rua Alfredo Allen, 4200-135 Porto, Portugal, E-mail: cflima@fpce.up.pt or c.lima@ucl.ac.uk.

DOI:10.1523/JNEUROSCI.2988-14.2014

Copyright © 2014 the authors 0270-6474/14/3412950-03\$15.00/0

physical differences and activity in bilateral superior temporal gyri (STG), left mid-temporal gyrus, right mid-cingulum, right precuneus, and right amygdala cannot be taken as evidence that these regions selectively encode the vocalizations' low-level features.

To examine higher-order processes, Bestelmeyer et al. (2014) capitalized on categorical perception effects, previously reported for speech prosody (Laukka, 2005) and nonverbal vocalizations (Bestelmeyer et al., 2010). Categorical perception occurs when continuous physical changes of stimuli do not result in continuous perceptual changes, but instead are perceived to fall into discrete categories: equal-sized physical changes that cause small perceptual shifts when occurring within a category will cause larger shifts in perception when they occur closer to the category boundary. After explaining variance accounted for by linear physical differences, Bestelmeyer et al. (2014) examined neural adaptation effects reflecting predominantly categorization shifts that occurred for stimuli near the middle of the morphing continuum. Quadratic associations were found in a network including bilateral anterior insulae, precuneus, mid-cingulum, left supplementary motor area (SMA), bilateral precentral gyri, left inferior frontal gyrus (IFG), left superior frontal gyrus, right STG, and bilateral medial superior frontal gyri. These quadratic associations indicate that hemodynamic responses to a given vocalization were similar when the preceding one elicited either identical or maximally different proportions of fear categorizations (i.e., perceptually identical and maximally different consecutive trials produced similar hemodynamic responses in these regions); and they were highest (or lowest, in the case of negative associations) for consecutive trials eliciting intermediately different perceptual changes (41–60%).

Finding quadratic, as opposed to linear, trends was unexpected, and no mechanistic account was offered in the paper. We argue that the quadratic trends may index emotional ambiguity and associated categorization/task difficulty. For identical and maximally different consecutive trials, the two vocalizations express similarly clear—or similarly ambiguous—categories, i.e., no differences exist in categorization difficulty. For intermediately different consecutive trials, however, is it more likely than the two vocalizations differ in ambiguity (e.g., a highly ambiguous vocalization following a clear one). Thus, a parsimonious interpretation of

these findings is that this network responds to differential difficulty in the forced categorization of vocalizations as “angry” or “fearful.” Longer response latencies for morphs in the middle of the continuum than for those closer to the extremes corroborate the potential for task difficulty effects.

A role for mentalizing and sensorimotor systems

Among the systems thought to mediate higher-order processing, two are particularly interesting in the context of the existing literature, although they are not discussed in the paper. Like Bestelmeyer et al. (2014), McGettigan et al. (2013) report medial prefrontal responses in a study on authenticity perception in laughter. During passive listening, these sites responded more strongly to voluntary social-type laughter than to genuine amusement laughter. Additionally, the magnitude of responses predicted performance in an off-line authenticity detection task. This was taken to reflect the automatic interpretation of intentions associated with social laughter, which is arguably more ambiguous than amusement laughter. Previously, medial prefrontal cortex was linked to person perception and attribution of mental states (mentalizing; Amodio and Frith, 2006). Because participants did not perform any task, McGettigan et al.'s (2013) findings cannot be attributed to task difficulty or motor responses. Thus, the new results of Bestelmeyer et al. (2014), combined with previous findings, suggest that mentalizing systems are part of the network involved in perceiving vocalizations, highlighting the social nature of these signals (Brück et al., 2011). Mentalizing may provide a mechanism for (1) resolving ambiguity—stimulus-driven ambiguity in the case of morphed vocalizations, and socially driven ambiguity in the case of different types of laughter; and (2) making socio-emotional inferences of varying complexity: basic emotion categories (Bestelmeyer et al., 2014) and nuanced within-category distinctions (McGettigan et al., 2013).

Another interesting result Bestelmeyer et al. (2014) report is modulation in sensorimotor systems including SMA, precentral gyrus, and IFG. Warren et al. (2006) reported activations in the same systems during passive listening of nonverbal vocalizations and also during the execution of facial movement, i.e., they are part of an auditory–motor mirror network. This suggests that a perception-to-action pathway contributes to under-

standing vocalizations. McGettigan et al. (2013) provided support for this hypothesis by showing a behavioral benefit (enhanced accuracy in authenticity detection) associated with functional responses within sensorimotor sites, possibly reflecting simulation of actions involved in the production of emotional sounds. Considering Bestelmeyer et al.'s (2014) task, one can speculate that sensorimotor responses reflect participants' effort to categorize ambiguous stimuli.

Future directions

These insights into vocal emotional communication raise many questions. It would be interesting, for instance, to examine the coding of low-level features by exploring within-category variability; e.g., laughter is highly variable but well recognized, allowing one to look at acoustic variability while controlling for higher-order aspects (e.g., emotion category; arousal). Bestelmeyer et al.'s (2014) study, along with previous findings, provide persuasive evidence for mentalizing and sensorimotor processes in the evaluation of vocalizations. Their role is nevertheless not considered to be central—or considered at all—in theoretical accounts of vocal affect (Schirmer and Kotz, 2006; Brück et al., 2011). Future research could explore whether these systems are more strongly recruited when the ambiguity of the stimuli poses more challenges in social cognition, and whether sensorimotor systems are more important for nonverbal vocalizations than for prosody because of their more direct link to action and to evolutionarily ancient forms of communication.

The stimuli used by Bestelmeyer et al. (2014) consist of the vowel /a/ as produced to express anger and fear, which is a step forward in expanding previous prosody-based research in vocal communication. Future work will benefit from further investigating different kinds of non-speech vocal stimuli—including more naturalistic ones (Scott et al., 1997; McGettigan et al., 2013)—for which relatively little research exists to date. Similarly to speech prosody, the perceived emotions in nonverbal vocalizations can be predicted from acoustic features (Sauter et al., 2010). Nonverbal vocalizations are yet distinct from prosody in important ways, not being constrained by verbal information, involving distinct production mechanisms, and arguably reflecting a primitive form of communication shared with other animal species (Juslin and Laukka, 2003).

Other aspects will need to be addressed as well, such as the bias toward negative

expressions in emotion research. Greater emphasis on positive vocalizations that are frequently encountered in everyday life may yield more insight. Including tasks and stimuli covering fine-grained aspects of emotion processing (e.g., continuous rating tasks; within-category perceptual distinctions) may also contribute to a nuanced and ecological view on vocal communication. Furthermore, current models need to be validated and expanded to account for individual differences, for instance due to development across the life span, personality, or neuropsychiatric conditions. Nonverbal vocalizations, however simple they might appear, hold promise as a valuable tool for probing the neurocognitive organization of both basic and sophisticated aspects of human social-communicative functioning.

References

- Aguirre GK (2007) Continuous carry-over designs for fMRI. *Neuroimage* 35:1480–1494. [CrossRef Medline](#)
- Amodio DM, Frith CD (2006) Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci* 7:268–277. [CrossRef Medline](#)
- Bestelmeyer PE, Rouger J, DeBruine LM, Belin P (2010) Auditory adaptation in vocal affect perception. *Cognition* 117:217–223. [CrossRef Medline](#)
- Bestelmeyer PE, Maurage P, Rouger J, Latinus M, Belin P (2014) Adaptation to vocal expressions reveals multistep perception of auditory emotion. *J Neurosci* 34:8098–8105. [CrossRef Medline](#)
- Brück C, Kreifelts B, Wildgruber D (2011) Emotional voices in context: a neurobiological model of multimodal affective information processing. *Phys Life Rev* 8:383–403. [CrossRef Medline](#)
- Darwin C (1872/2009) *The expression of the emotions in man and animals*. London UK: Penguin Books.
- Juslin PN, Laukka P (2003) Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol Bull* 129:770–814. [CrossRef Medline](#)
- Laukka P (2005) Categorical perception of vocal emotion expressions. *Emotion* 5:277–295. [CrossRef Medline](#)
- McGettigan C, Walsh E, Jessop R, Agnew ZK, Sauter DA, Warren JE, Scott SK (2013) Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cereb Cortex*. Advance online publication. Retrieved Aug. 22, 2013. doi:10.1093/cercor/bht227. [CrossRef Medline](#)
- Sauter DA, Eisner F, Calder AJ, Scott SK (2010) Perceptual cues in nonverbal vocal expressions of emotion. *Q J Exp Psychol (Hove)* 63:2251–2272. [CrossRef Medline](#)
- Schirmer A, Kotz SA (2006) Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn Sci* 10:24–30. [CrossRef Medline](#)
- Scott SK, Young AW, Calder AJ, Hellawell DJ, Aggleton JP, Johnson M (1997) Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature* 385:254–257. [CrossRef Medline](#)
- Warren JE, Sauter DA, Eisner F, Wiland J, Dresner MA, Wise RJ, Rosen S, Scott SK (2006) Positive emotions preferentially engage an auditory-motor “mirror” system. *J Neurosci* 26:13067–13075. [CrossRef Medline](#)