

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.

Distributed Sensitivity to Conspecific Vocalizations and Implications for the Auditory Dual Stream Hypothesis

Jennifer K. Bizley* and Kerry M. M. Walker*

Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford OX1 3PT, United Kingdom
Review of Recanzone (<http://www.jneurosci.org/cgi/content/full/28/49/13184>)

The dual stream hypothesis posits that auditory cortex contains two parallel and hierarchical processing streams that are independently specialized for sound localization and identification. Anatomical studies have demonstrated a putative structural basis for these pathways, beginning as early as the auditory cortical belt (Romanski et al., 1999). The neurophysiological evidence for a caudal sound localization stream has been compelling (Woods et al., 2006). However, the proposed specialization of more rostral regions for auditory object identification, specifically vocalization identification, is currently less clear. Although one study has demonstrated that single neurons in the rostral auditory cortex of primates respond more selectively to conspecific vocalizations than caudal regions (Tian et al., 2001), others emphasize that neurons in separate auditory cortical fields show only subtle differences in their response properties, including their selectivity for conspecific vocalizations. Moreover, imaging studies have demonstrated that vocalizations invoke robust activity across much of auditory cortex, and specializa-

tion for conspecific vocalizations is observed only in anterior regions beyond auditory cortex (Petkov et al., 2008). Thus, the precise specialization of the rostral auditory stream remains elusive, as does the existence of a dedicated “conspecific vocalization” region at the level of auditory cortex. These issues were recently addressed by Recanzone (2008) in the *Journal of Neuroscience*.

Recanzone (2008) presented four conspecific vocalizations to awake monkeys while recording the responses of single neurons across five auditory cortical fields: two core [primary auditory cortex (A1) and rostral field (R)] and three belt [caudolateral (CL), caudomedial (CM), and middle lateral (ML)] regions (Fig. 1). Single neuron selectivity for these calls was assessed using three metrics. First, a monkey call preference index (MCI_{50}) measured how many of the calls evoked a firing rate that matched or exceeded 50% of the firing rate in response to the preferred call. Because four calls were presented, the MCI_{50} took the form of an integer value, from 1 (most selective) to 4 (least selective), with a value of 1 indicating that only a single call evoked the maximal firing rate. Second, the MCI_t selectivity index was calculated as the number of vocalizations that elicited a firing rate that was significantly different from the response to the preferred call. The MCI_t also took the form of an integer value from 1 (most selective) to 4 (least selective). Finally, a pattern discrimination algorithm based on the Euclidean distances between

spike trains was used to measure the discriminability of neurons' temporal spiking patterns in response to different vocalizations.

One of the most striking findings in this study was that neurons across all five cortical fields showed a similar degree of selectivity for monkey vocalizations, as measured using either the MCI_t or the MCI_{50} . Approximately equal proportions of neurons (~25%) in each cortical field had an MCI_{50} of 1 for monkey calls presented in their natural state [Recanzone (2008), their Fig. 3A (<http://www.jneurosci.org/cgi/content/full/28/49/13184/F3>)]. A larger proportion of neurons (~40%) had an MCI_t of 1 across all five fields [Recanzone (2008), their Fig. 4A (<http://www.jneurosci.org/cgi/content/full/28/49/13184/F4>)], indicating that the spike rate in response to the preferred call was statistically distinguishable from the response to the other three calls, whereas only a small percentage of neurons (10–20%) were unselective for call identity. Overall, neither call index demonstrated a clear difference in the degree of vocalization selectivity across cortical fields, although this conclusion would have been strengthened by the use of statistical analysis to test this equivalence across regions.

To further test whether these cortical fields may be specialized for identifying conspecific vocalizations, the same four calls were presented time-reversed. The author reasons, like others before him, that by reversing a conspecific vocalization, the spectral content and complexity is maintained while the behavioral rele-

Received Feb. 5, 2009; revised Feb. 5, 2009; accepted Feb. 5, 2009.

J.K.B. was supported by Biotechnology and Biological Sciences Research Council Grant BB/D009758/1, and K.M.M.W. was supported by Wellcome Trust Grant 076508/Z/05/Z.

*J.K.B. and K.M.M.W. contributed equally to this work.

Correspondence should be addressed to either Dr. Jennifer K. Bizley or Dr. Kerry M. M. Walker, Department of Physiology, Anatomy and Genetics, Sherrington Building, University of Oxford, Oxford OX1 3PT, UK. E-mail: jennifer.bizley@dpag.ox.ac.uk or kerry.walker@dpag.ox.ac.uk.

DOI:10.1523/JNEUROSCI.6035-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/293011-03\$15.00/0

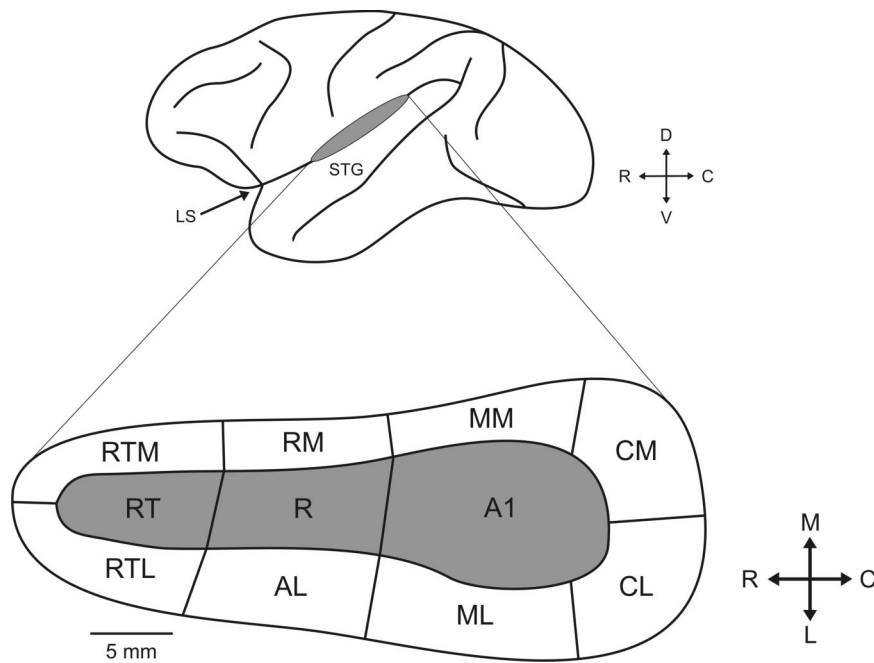


Figure 1. Schematic of the lateral view of the macaque brain (above), with the location of auditory cortex within the lateral sulcus indicated by a gray ellipse. The lower inset is a schematic identifying the three core (gray) and eight belt (white) regions. LS, Lateral sulcus; STG, superior temporal gyrus; RT, rostrotemporal field; AL, anterolateral area; RTL, lateral rostrotemporal area; RTM, medial rostrotemporal area; RM, rostromedial area; MM, middle medial area; D, dorsal; V, ventral; M, medial; L, lateral.

vance of the sound is disrupted. Both spike count measures failed to reveal even one neuron whose responses were selective for all four (or any three) of the natural calls over their time-reversed counterparts. Hence, there was no evidence of generic vocalization specificity within these fields. Nevertheless, a small proportion (<15%) of neurons were selective for the temporal direction of at least one call [Recanzone (2008), their Table 4 (<http://www.jneurosci.org/cgi/content/full/28/49/13184/T4>)]. Contrary to the prediction of rostral vocalization selectivity, only neurons in field CL fired at a significantly higher overall rate in response to natural calls compared with reversed calls.

Previous studies have demonstrated that a discriminator based on the temporal spiking patterns of auditory cortical neurons can distinguish between natural and time-reversed vocalizations, even where spike-rate-based discriminators fail (Schnupp et al., 2006). Along similar lines, Recanzone (2008) used the linear pattern discriminator developed by Schnupp et al. (2006) to show that the complete set of natural and time-reversed calls could be classified with ~90% accuracy based on the temporal response patterns of single neurons in all cortical fields.

Recanzone (2008) found that when relatively wide time bins (>25 ms) were used to decode neural responses, neurons

in field R supported more accurate discrimination of reversed calls than neurons in the other four cortical fields. In fact, in field R, but not in the other cortical fields examined, the linear pattern discriminator could classify neural responses to different reversed vocalizations better than the responses to the set of natural vocalizations at all resolutions beyond 25 ms, and this effect strengthened when wider time bins were used [Recanzone (2008), their Fig. 7D (<http://www.jneurosci.org/cgi/content/full/28/49/13184/F7>)]. Neurons in field R have been shown to have slower temporal dynamics than those in A1 (Bendor and Wang, 2008), which could explain why only neural responses in this field continued to support discrimination of reversed vocal calls at wider temporal integration windows. However, when responses were decoded with bin widths of <25 ms, discrimination of both forwards and reversed vocalizations was equivalent in neurons across all cortical fields.

In summary, none of the three selectivity indices used by Recanzone (2008) provided evidence of a regional specialization for vocalization processing within these auditory cortical subdivisions. This homogeneity of vocalization sensitivity across cortical fields is clearly in contrast to a strict interpretation of the dual stream hypothesis, which would predict that neurons in fields R and ML would be more

vocalization-selective than those in CM or CL. Although previous investigators have demonstrated that neurons in the even further rostral field, AL, tended to be more selective for vocalizations, they failed to demonstrate a difference in the vocalization selectivity of fields ML and CL (Tian et al., 2001). Thus, vocalization specialization may only become evident in fields rostral to those studied by Recanzone (2008). Indeed, a very recent study has identified a potentially call-selective auditory area in the macaque, located in the caudal insular cortex (Remedios et al., 2009). Recanzone (2008) acknowledges that the set of four calls he used may be insufficient to demonstrate call selectivity, because previous studies have often used a more diverse call set. Additionally, the stimulus-specific responses described by Tian et al. (2001) may not in fact reflect vocalization specificity, but a more general feature sensitivity that can be demonstrated by presenting both vocal and non-vocal complex sounds.

We would add a further point for consideration in interpreting the results of these studies. In the experiment by Recanzone (2008), the responses of auditory cortical neurons were recorded from monkeys while they were engaged in a listening task that required them to report a change in location of the sound source while deliberately ignoring call identity changes. This raises the possibility that neural tuning to vocalization identity was suppressed during this task and that a greater vocalization selectivity might be observed under different behavioral conditions. Interestingly, in a study that demonstrated the spatial sensitivity of neurons in caudal areas in these same animals, the experimental paradigm allowed animals to use sound-source location as a predictor of reward (Woods et al., 2006).

Although studies of the relative sensitivity to vocal call types (or features) can offer useful insights into cortical organization in the absence of claims regarding the stimulus specificity of those neural responses, the question of feature selectivity in auditory streams could be further examined by demonstrating how independently these parameters are encoded in the same neurons. That is, do any of these neurons represent vocalization identity in a manner which is independent of the spatial location of the caller? Evidence of this type of feature selectivity among neurons in a particular cortical field would strongly suggest a specialization for vocalization recognition.

Reversed vocalization stimuli have

been frequently used to assess specialization for conspecific calls, but it is important to note that although this procedure preserves the overall spectral energy of the stimulus, it severely alters a number of temporal features. For instance, a stimulus with a sharp onset and a gradual offset will be changed to one with a slower rise time and an abrupt endpoint. Neurons in auditory cortex are exquisitely sensitive to the precise temporal properties of sounds, notably to the rate of increase in sound pressure at the sound onset (Heil, 1997), and the direction and rate of frequency modulations (Tian and Rauschecker, 2004). Consequentially, the temporal response properties of a given neuron will determine whether it produces distinguishable responses to natural and time-reversed calls. Quite logically, Recanzone (2008) points out that the superior discrimination of reversed calls in field R is unlikely to reflect specialized processing for these sounds because reversed vocalizations do not naturally occur. However, had neurons in a particular field been shown to encode forwards vocalizations more accurately than reversed ones, caution should be similarly exercised before such a result is interpreted as evidence for a cortical region that is specialized for conspecific vocalization processing. Rather, a more empirically rigorous approach that rules out tuning to basic

acoustical processes should be undertaken. One very promising tool that could be used in such an analysis are the “virtual vocalizations” developed by DiMattina and Wang (2006), which allow multiple properties of a vocalization to be manipulated systematically both within and outside of the naturalistic range. A region that is specialized for the identification of conspecific vocalizations would be expected to exhibit selectivity for such behaviorally relevant acoustic features, while remaining relatively invariant to features that are ecologically implausible or irrelevant to call identification.

The question of dual processing streams in auditory cortex has inspired debate and motivated experimentation over the past decade. The functional role(s) of the rostral stream remains unclear, but the work of Recanzone (2008) reminds us that neural specificity for conspecific vocalizations might be unlikely at the level of core and belt auditory cortex. Instead, these regions may be specialized for processing more general features of complex sounds, the nature of which are yet to be discovered.

References

- Bendor D, Wang X (2008) Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *J Neurophysiol* 100:888–906.
- DiMattina C, Wang X (2006) Virtual vocalization stimuli for investigating neural representations of species-specific vocalizations. *J Neurophysiol* 95:1244–1262.
- Heil P (1997) Auditory cortical onset responses revisited. I. First-spike timing. *J Neurophysiol* 77:2616–2641.
- Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK (2008) A voice region in the monkey brain. *Nat Neurosci* 11:367–374.
- Recanzone GH (2008) Representation of conspecific vocalizations in the core and belt areas of the auditory cortex in the alert macaque monkey. *J Neurosci* 28:13184–13193.
- Remedios R, Logothetis NK, Kayser C (2009) An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. *J Neurosci* 29:1034–1045.
- Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci* 2:1131–1136.
- Schnupp JW, Hall TM, Kokelaar RF, Ahmed B (2006) Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J Neurosci* 26:4785–4795.
- Tian B, Rauschecker JP (2004) Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol* 92:2993–3013.
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292:290–293.
- Woods TM, Lopez SE, Long JH, Rahman JE, Recanzone GH (2006) Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. *J Neurophysiol* 96:3323–3337.