



Neural mechanisms of auditory categorization: from across brain areas to within local microcircuits

Joji Tsunada^{1*} and Yale E. Cohen^{1,2,3}

¹ Department of Otorhinolaryngology-Head and Neck Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

² Department of Neuroscience, University of Pennsylvania, Philadelphia, PA, USA

³ Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA

Edited by:

Einat Liebenthal, Medical College of Wisconsin, USA

Reviewed by:

Amy Poremba, University of Iowa, USA

Rutvik Desai, University of South Carolina, USA

*Correspondence:

Joji Tsunada, Department of Otorhinolaryngology-Head and Neck Surgery, Perelman School of Medicine, University of Pennsylvania, 3400 Spruce Street, 5 Ravdin, Philadelphia, PA 19104, USA

Office: Perelman School of Medicine, University of Pennsylvania, John Morgan Building B50, 3620 Hamilton Walk, Philadelphia, PA 19104-6055, USA
e-mail: tsunada@mail.med.upenn.edu

Categorization enables listeners to efficiently encode and respond to auditory stimuli. Behavioral evidence for auditory categorization has been well documented across a broad range of human and non-human animal species. Moreover, neural correlates of auditory categorization have been documented in a variety of different brain regions in the ventral auditory pathway, which is thought to underlie auditory-object processing and auditory perception. Here, we review and discuss how neural representations of auditory categories are transformed across different scales of neural organization in the ventral auditory pathway: from across different brain areas to within local microcircuits. We propose different neural transformations across different scales of neural organization in auditory categorization. Along the ascending auditory system in the ventral pathway, there is a progression in the encoding of categories from simple acoustic categories to categories for abstract information. On the other hand, in local microcircuits, different classes of neurons differentially compute categorical information.

Keywords: auditory category, ventral auditory pathway, speech sound, vocalization, pyramidal neuron, interneuron

INTRODUCTION

Auditory categorization is a computational process in which sounds are classified and grouped based on their acoustic features and other types of information (e.g., semantic knowledge about the sounds). For example, when we hear the word “Hello” from different speakers, we can categorize the gender of each speaker based on the pitch of the speaker’s voice. On the other hand, in order to analyze the linguistic content transmitted by speech sounds, we can ignore the unique pitch, timbre etc. of each speaker and categorize the sound into the distinct word category “Hello.” Thus, auditory categorization enables humans and non-human animals to extract, manipulate, and efficiently respond to sounds (Miller et al., 2002, 2003; Russ et al., 2007; Freedman and Miller, 2008; Miller and Cohen, 2010).

A specific type of categorization is called “categorical perception” (Liberman et al., 1967; Kuhl and Miller, 1975, 1978; Kuhl and Padden, 1982, 1983; Kluender et al., 1987; Pastore et al., 1990; Lotto et al., 1997; Sinnott and Brown, 1997; Holt and Lotto, 2010). The primary characteristic of categorical perception is that the perception of a sound does not smoothly vary with changes in its acoustic features. That is, in certain situations, small changes in the physical properties of an acoustic stimulus can cause large changes in a listener’s perception of a sound. In other situations, large changes can cause no change in perception. The stimuli, which cause these large changes in perception, straddle the boundary between categories. For example,

when we hear a continuum of smoothly varying speech sounds (i.e., a continuum of morphed stimuli between the phoneme prototypes “ba” and “da”), we experience a discrete change in perception. Specifically, a small change in the features of a sound near the middle of this continuum (i.e., at the category boundary between a listener’s perception of “ba” and “da”) will cause a large change in a listener’s perceptual report. In contrast, when that same small change occurs at one of the ends of the continuum, there is little effect on the listener’s report.

Even though some perceptual categories have sharp boundaries, the locations of the boundary are somewhat malleable. For instance, the perception of a phoneme can be influenced by the phonemes that come before it. When morphed stimuli, which are made from the prototypes “da” and “ga,” are preceded by presentations of “al” or “ar,” the perceptual boundary between the two prototypes shifts (Mann, 1980). Specifically, listeners’ reports are biased toward reporting the morphed stimuli as “da” when it is preceded by “ar.” When this morphed stimulus is preceded by “al,” listeners are biased toward reporting the morphed stimulus as “ga.”

Categories are not only formed based on the perceptual features of stimuli but also on more “abstract” types of information. An abstract category is one in which a group of arbitrary stimuli are linked together as a category based on some shared features, a common functional characteristic, semantic information,

or acquired knowledge. For instance, a combination of physical characteristics and knowledge about their reproductive processes puts dogs, cats, and killer whales into one category (“mammals”), but birds into a separate category. However, if we use different criteria to form a category of “pets,” dogs, cats, and birds would be members of this “pet” category but not killer whales.

Behavioral responses to auditory communication signals (i.e., species-specific vocalizations) also provide evidence for abstract categorization. One example is the categorization of food-related species-specific vocalizations by rhesus monkeys (Hauser and Marler, 1993a,b; Hauser, 1998; Gifford et al., 2003). In rhesus monkeys, a vocalization called a “harmonic arch” transmits information about the discovery of rare, high-quality food. A different vocalization called a “warble” also transmits the same type of information: the discovery of rare, high-quality food. Importantly, whereas both harmonic arches and warbles transmit the same type of information, they have distinct spectrotemporal properties. Nevertheless, rhesus monkeys’ responses to those vocalizations indicate that monkeys categorize these two calls based on their transmitted information and not their acoustic features. In another example, Diana monkeys form abstract-categorical representations for predator-specific alarm calls independent of the species generating the signal. Diana monkeys categorize and respond similarly to alarm calls that signify the presence of a leopard, regardless of whether the alarm calls are elicited from a Diana monkey or a crested guinea fowl (Zuberbühler and Seyfarth, 1997; Zuberbühler, 2000a,b). Similarly, Diana monkeys show similar categorical-responses to eagle alarm calls that can be elicited from other Diana monkeys or from putty-nose monkeys (Eckardt and Zuberbühler, 2004).

In order to better understand the mechanisms that underlie auditory categorization, it is essential to examine how neural representations of auditory categories are formed and transformed across different scales of neural organization: from across different brain areas to within local microcircuits. In this review, we discuss the representation of auditory categories in different cortical regions of the ventral auditory pathway; the hierarchical processing of categorical information along the ventral pathway; and the differential role that excitatory pyramidal neurons and inhibitory interneurons (i.e., different neuron classes) contribute to these categorical computations.

The ventral pathway is targeted because neural computations in this pathway are thought to underlie sound perception, which is critically related to auditory categorization and auditory scene analysis (Rauschecker and Scott, 2009; Romanski and Averbeck, 2009; Bizley and Cohen, 2013). The ventral auditory pathway begins in the core auditory cortex (in particular, the primary auditory cortex and the rostral field R) and continues into the anterolateral and middle-lateral belt regions. These belt regions then project either directly or indirectly to the ventral prefrontal cortex (**Figure 1**) (Hackett et al., 1998; Rauschecker, 1998; Kaas and Hackett, 1999, 2000; Kaas et al., 1999; Romanski et al., 1999a,b; Rauschecker and Tian, 2000; Rauschecker and Scott, 2009; Romanski and Averbeck, 2009; Recanzone and Cohen, 2010; Bizley and Cohen, 2013).

NEURAL TRANSFORMATIONS ACROSS CORTICAL AREAS IN THE VENTRAL AUDITORY PATHWAY

In this section, we discuss how auditory categories are processed in the ventral auditory pathway. More specifically, we review the representation of auditory categories across different regions in the ventral auditory pathway and then discuss the hierarchical processing of categorical information in the ventral auditory pathway.

Before we continue, it is important to define the concept of a “neural correlate of categorization.” One simple definition is the following: a neural response is “categorical” when the responses are invariant to the stimuli that belong to the same category. In practice, neuroimaging techniques define “categorical” responses as equivalent activations of distinct brain regions by within-category stimuli and the equivalent activation of different brain regions by stimulus exemplars from a second category (Binder et al., 2000; Altmann et al., 2007; Doehrmann et al., 2008; Leaver and Rauschecker, 2010). At the level of single neurons, a neuron is said to be “categorical” if its firing rate is invariant to different members of one category and if it has a second level of (invariant) responsiveness to stimulus exemplars from a second category (Freedman et al., 2001; Tsunada et al., 2011). The specific mechanisms that underlie the creation of category sensitive neurons are not known. However, presumably, they rely on the computations that mediate stimulus invariance in neural selectivity and perception (Logothetis and Sheinberg, 1996; Holt and Lotto, 2010; Dicarolo et al., 2012). Moreover, because animals can form a wide range of categories based on individual experiences, a degree of learning and plasticity must be involved in the creation of *de-novo* category selective responses (Freedman et al., 2001; Freedman and Assad, 2006). Indeed, when monkeys were trained to categorize stimuli with different category boundaries, boundaries for categorical responses in some brain areas (e.g., the prefrontal and parietal cortices) also changed (Freedman et al., 2001; Freedman and Assad, 2006).

HOW DO DIFFERENT CORTICAL AREAS IN THE VENTRAL AUDITORY PATHWAY SIMILARLY OR DIFFERENTIALLY REPRESENT CATEGORICAL INFORMATION?

It is well known that neurons become increasingly sensitive to more complex stimuli and abstract information between the beginning stages of the ventral auditory pathway (i.e., the core) and the latter stages (e.g., the ventral prefrontal cortex). For example, neurons in the core auditory cortex are more sharply tuned for tone bursts than neurons in the lateral belt (Rauschecker et al., 1995), whereas lateral-belt neurons are more sensitive to the spectrotemporal properties of complex sounds, such as vocalizations (Rauschecker et al., 1995; Tian and Rauschecker, 2004). Furthermore, beyond the auditory cortex, the ventral prefrontal cortex not only encodes complex sounds (Averbeck and Romanski, 2004; Cohen et al., 2007; Russ et al., 2008a; Miller and Cohen, 2010) but also has a critical role for attention and memory-related cognitive functions (e.g., memory retrieval) which are critical for abstract categorization (Goldman-Rakic, 1995; Miller, 2000; Miller and Cohen, 2001; Miller et al., 2002, 2003; Gold and Shadlen, 2007; Osada et al., 2008; Cohen et al., 2009; Plakke et al., 2013a,b,c; Poremba et al., 2013).

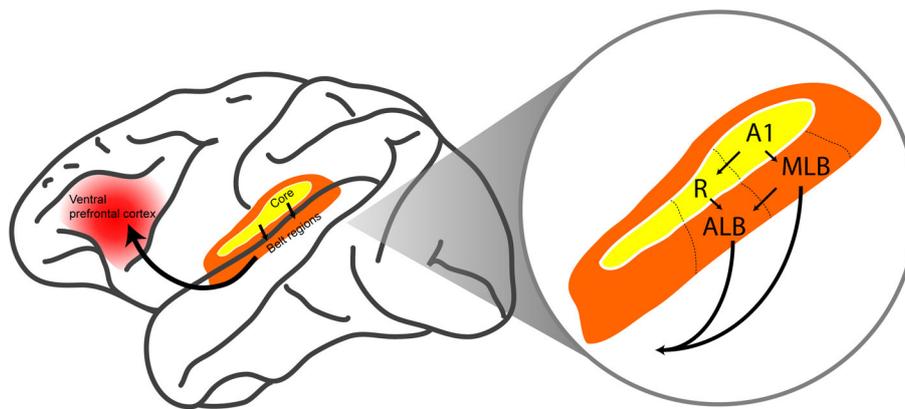


FIGURE 1 | The ventral auditory pathway in the monkey brain.

The ventral auditory pathway begins in core auditory cortex (in particular, the primary auditory cortex A1 and the rostral field R). The pathway continues into the middle-lateral (MLB) and

anterolateral (ALB) belt regions, which project directly and indirectly to the ventral prefrontal cortex. Arrows indicate feedforward projections. The figure is modified, with permission, from Hackett et al. (1998) and Romanski et al. (1999a).

These observations are consistent with the idea that there is a progression of category-information processing along the ventral auditory pathway: brain regions become increasingly sensitive to more complex types of categories. More specifically, it appears that neurons in core auditory cortex may encode categories for simple sounds, whereas neurons in the belt regions and the ventral prefrontal cortex may encode categories for more complex sounds and abstract information.

Indeed, neural correlates of auditory categorization can be seen in the core auditory cortex for simple frequency contours (Ohl et al., 2001; Selezneva et al., 2006). For example, in a study by Selezneva and colleagues, monkeys categorized the direction of a frequency contour of tone-burst sequences as either “increasing” or “decreasing” while neural activity was recorded from the primary auditory cortex. Selezneva et al. found that these core neurons encoded the sequence direction independent of its specific frequency content: that is, a core neuron responded similarly to a decreasing sequence from 1 to 0.5 kHz as it did to a decreasing sequence from 6 to 3 kHz. In a second study, Ohl et al. demonstrated that categorical representations need not be represented in the firing rates of single neurons but, instead, can be encoded in the dynamic firing patterns of a neural population. Thus, even in the earliest stage of the ventral auditory pathway, there is evidence for neural categorization.

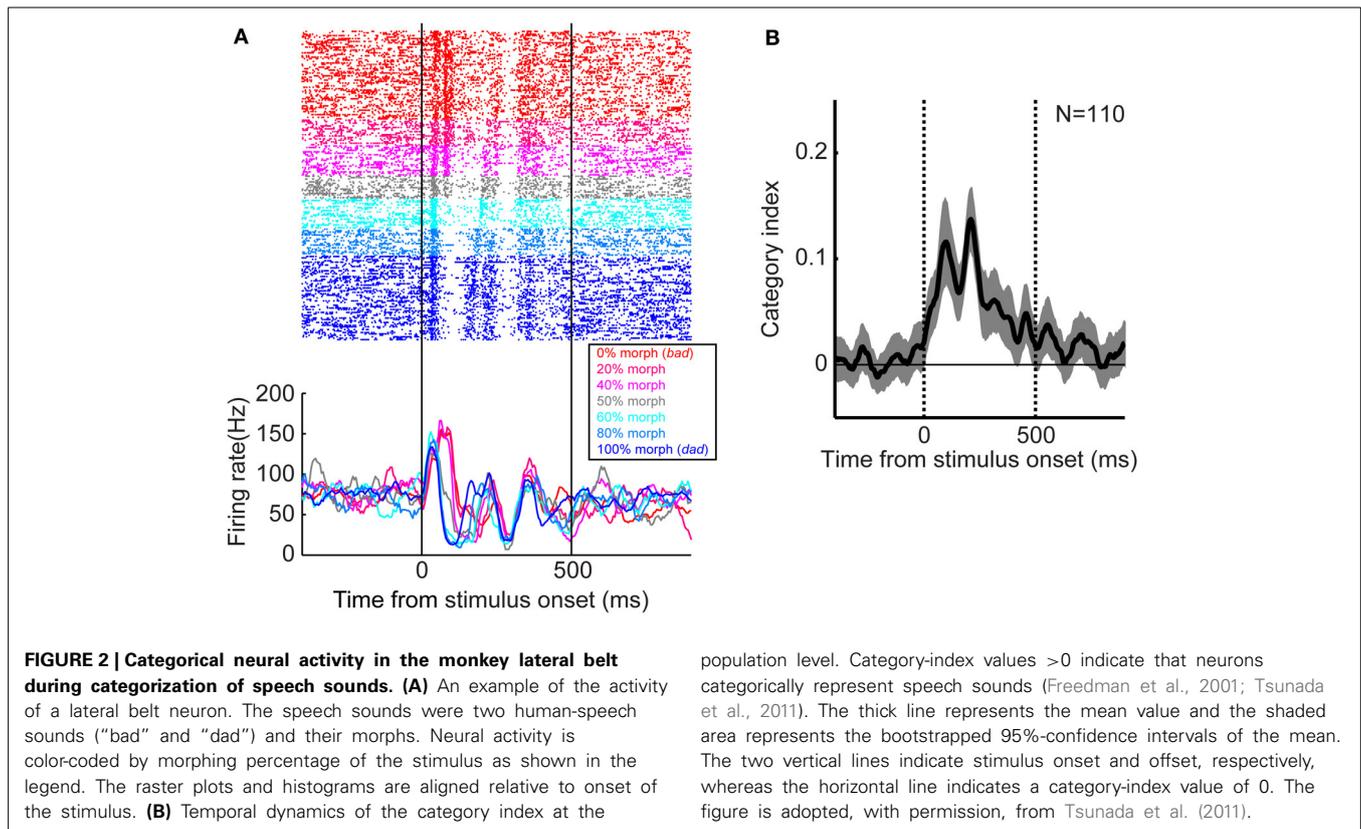
Although the core auditory cortex processes categorical information for simple auditory stimuli (e.g., the direction of frequency changes of pure tones), studies using more complex sounds, such as human-speech sounds, have shown that core neurons primarily encode the acoustic features that compose these complex sounds but do not encode their category membership (Liebenthal et al., 2005; Steinschneider et al., 2005; Obleser et al., 2007; Engineer et al., 2008, 2013; Mesgarani et al., 2008, 2014; Nourski et al., 2009; Steinschneider, 2013). That is, the categorization of complex sounds requires not only analyses at the level of the acoustic feature but also subsequent computations that integrate the analyzed features into a perceptual representation, which is then subject to a categorization process. For example,

distributed and temporally dynamic neural responses in individual core neurons can represent different acoustic features of speech sounds (Schreiner, 1998; Steinschneider et al., 2003; Engineer et al., 2008; Mesgarani et al., 2008, 2014), but the categorization of the speech sounds requires classifying the activation pattern across the entire population of core neurons.

Categorical representations of speech sounds at the level of the single neuron or local populations of neurons appear to occur at the next stage of auditory processing in the ventral auditory pathway, the lateral-belt regions. Several recent studies have noted that neural activity in the monkey lateral-belt and human superior temporal gyrus encodes speech-sound categories (Chang et al., 2010; Steinschneider et al., 2011; Tsunada et al., 2011; Steinschneider, 2013). For example, our group found that, when monkeys categorized two prototypes of speech sounds (“bad” and “dad”) and their morphed versions, neural activity in the lateral belt discretely changed at the category boundary, suggesting that these neurons encoded the auditory category rather than smoothly varying acoustic features (**Figure 2**).

Human-neuroimaging studies have also found that the superior temporal sulcus is categorically activated by speech sounds, relative to other sounds (Binder et al., 2000; Leaver and Rauschecker, 2010). Specifically, the superior temporal sulcus was activated more by speech sounds than by frequency-modulated tones (Binder et al., 2000) or by other sounds including bird songs and animal vocalizations (Leaver and Rauschecker, 2010). Furthermore, activity in the superior temporal sulcus did not simply reflect the acoustic properties of speech sounds but, instead, represented the perception of speech (Mottonen et al., 2006; Desai et al., 2008).

Additionally, studies with other complex stimuli provide further evidence for the categorical encoding of complex sounds in the human non-primary auditory cortex, including superior temporal gyrus and sulcus, but not in the core auditory cortex (Altmann et al., 2007; Doehrmann et al., 2008; Leaver and Rauschecker, 2010). These studies found that complex sound categories were represented in spatially distinct and widely



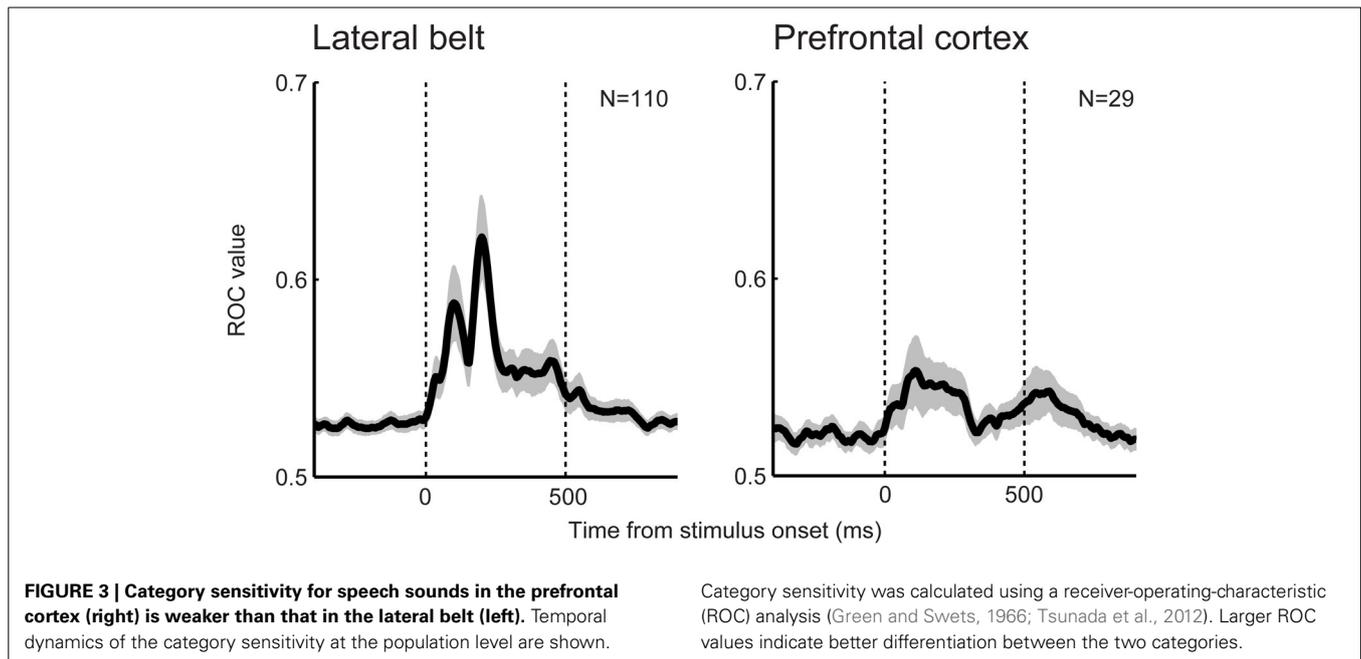
population level. Category-index values >0 indicate that neurons categorically represent speech sounds (Freedman et al., 2001; Tsunada et al., 2011). The thick line represents the mean value and the shaded area represents the bootstrapped 95%-confidence intervals of the mean. The two vertical lines indicate stimulus onset and offset, respectively, whereas the horizontal line indicates a category-index value of 0. The figure is adopted, with permission, from Tsunada et al. (2011).

distributed sub-regions within the superior temporal gyrus and sulcus (Oblaser et al., 2006, 2010; Engel et al., 2009; Staeren et al., 2009; Chang et al., 2010; Leaver and Rauschecker, 2010; Giordano et al., 2013). For example, distinct regions of the superior temporal gyrus and sulcus are selectively activated by musical-instrument sounds (Leaver and Rauschecker, 2010), tool sounds (Doehrmann et al., 2008), and human-speech sounds (Belin et al., 2000; Binder et al., 2000; Warren et al., 2006); whereas the anterior part of the superior temporal gyrus and sulcus is preferentially activated by the passive listening of conspecific vocalizations than other vocalizations (Fecteau et al., 2004). Similar findings for con-specific vocalizations have been obtained in the monkey auditory cortex (Petkov et al., 2008; Perrodin et al., 2011). Consistent with these findings, neuropsychological studies have shown that human patients with damage in the temporal cortex have deficits in voice recognition and discrimination (i.e., phonagnosia Van Lancker and Canter, 1982; Van Lancker et al., 1988; Goll et al., 2010). Thus, hierarchically higher regions in the auditory cortex encode complex-sound categories in spatially distinct (i.e., modular) and widely distributed sub-regions.

Moreover, recent studies posit that the sub-regions in the non-primary auditory cortex process categorical information in a hierarchical manner (Warren et al., 2006). A recent meta-analysis of human speech-processing studies suggests that a hierarchical organization of speech processing exists within the superior temporal gyrus: the middle superior temporal gyrus is sensitive to phonemes; anterior superior temporal gyrus to words; and the most anterior locations to short phrases (Dewitt and

Rauschecker, 2012; Rauschecker, 2012). Additionally, a different hierarchical processing of speech sounds in the superior temporal sulcus has also been articulated: the posterior superior temporal sulcus is preferentially sensitive for newly acquired sound categories, whereas the middle and anterior superior temporal sulci are more responsive to familiar sound categories (Liebenthal et al., 2005, 2010). Thus, within different areas of the non-primary auditory cortex, multiple and parallel processing may progress during auditory categorization.

Beyond the auditory cortex, do latter processing stages (e.g., the monkey ventral prefrontal cortex and human inferior frontal cortex) process categories for even more complex sounds? A re-examination of previous findings from our lab (Russ et al., 2008b; Tsunada et al., 2011) indicated important differences in neural categorization between the lateral belt and the ventral prefrontal cortex (**Figure 3**). We found that, at the population level, the category sensitivity for speech sounds in the prefrontal cortex was weaker than that in the lateral belt although neural activity in the prefrontal cortex transmitted a significant amount of categorical information. Consistent with this finding, a human-neuroimaging study also found that neural activity in the superior temporal gyrus is better correlated with a listener’s ability to discriminate between speech sounds than the activity in the inferior prefrontal cortex (Binder et al., 2004). Because complex sounds, including speech sounds, are substantially processed in the non-primary auditory cortex as discussed above, the prefrontal cortex may not represent, relative to the auditory cortex, a higher level of auditory perceptual-feature categorization.



Instead, the prefrontal cortex may be more sensitive to categories that are formed based on the abstract information that is transmitted by sounds. For example, the human inferior prefrontal cortex may encode categories for abstract information like emotional valence of a speaker's voice (Fecteau et al., 2005). Furthermore, human electroencephalography and neuroimaging studies have also revealed that the inferior prefrontal cortex plays a key role in the categorization of semantic information of multisensory stimuli (Werner and Noppeney, 2010; Joassin et al., 2011; Hu et al., 2012): Joassin et al. showed that the inferior prefrontal cortex contains multisensory category representations of gender that is derived from a speaker's voice and from visual images of a person's face.

Similarly, the monkey ventral prefrontal cortex encodes abstract categories. We have found that neurons in the ventral prefrontal cortex represent categories for food-related calls based on the transmitted information (e.g., high quality food vs. low quality food) (Gifford et al., 2005; Cohen et al., 2006). A more recent study found that neural activity in the monkey prefrontal cortex categorically represents the number of auditory stimuli (Nieder, 2012). Thus, along the ascending auditory system in the ventral auditory pathway, cortical areas encode categories for more complex stimuli and more abstract information.

NEURAL TRANSFORMATIONS WITHIN LOCAL MICROCIRCUITS

In this section, we discuss how the categorical information represented in each cortical area of the ventral auditory pathway is computed within local microcircuits. First, we briefly review the cortical microcircuit. Next, we focus on the role that two main cell classes of neurons in cortical microcircuits (i.e., excitatory pyramidal neurons and inhibitory interneurons) and discuss how different classes of neurons process categorical information.

HOW DO DIFFERENT CLASSES OF NEURONS IN LOCAL MICROCIRCUITS PROCESS CATEGORICAL INFORMATION?

A cortical microcircuit can be defined as a functional unit that processes inputs and generates outputs by dynamic and local interactions of excitatory pyramidal neurons and inhibitory interneurons (Merchant et al., 2012). Consequently, pyramidal neurons and interneurons are considered to be the main elements of microcircuits. Pyramidal neurons, which consist ~70–90% of cortical neurons, provide excitatory-outputs locally (i.e., within a cortical area) and across brain areas (Markham et al., 2004). On the other hand, interneurons, which consist small portion of cortical neurons (~10–30%), provide mainly inhibitory-outputs to surrounding pyramidal neurons and other interneurons (Markham et al., 2004).

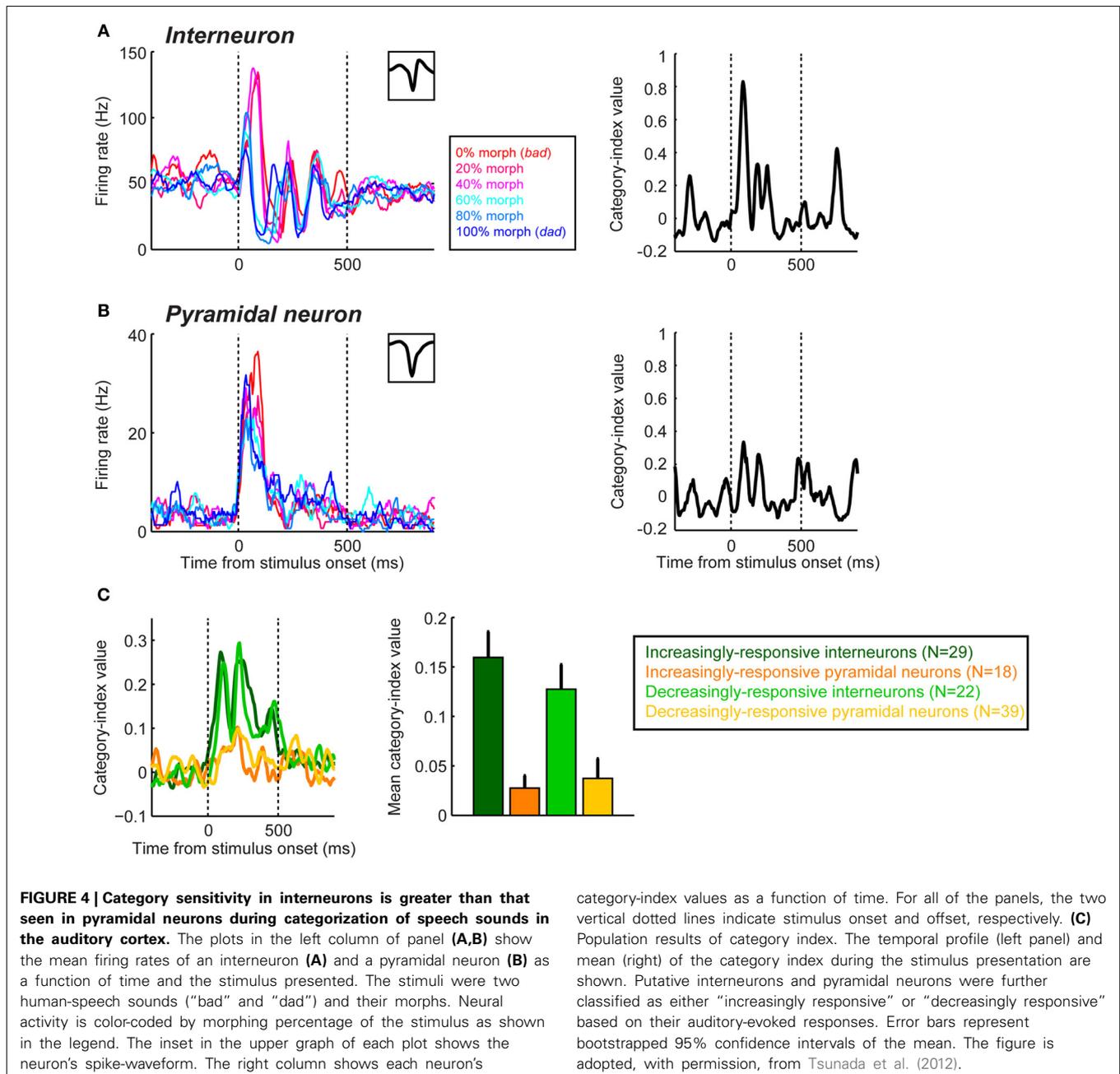
From a physiological perspective, pyramidal neurons and interneurons can be classified based on the waveform of their action potentials (Mountcastle et al., 1969; McCormick et al., 1985; Kawaguchi and Kubota, 1993, 1997; Kawaguchi and Kondo, 2002; Markham et al., 2004; González-Burgos et al., 2005). More specifically, the waveforms of pyramidal neurons tend to be broader and slower than those seen in the most interneurons. Using this classification, several extracellular-recording studies have been able to elucidate roles of pyramidal neurons and interneurons for visual working memory in the prefrontal cortex (Wilson et al., 1994; Rao et al., 1999; Constantinidis and Goldman-Rakic, 2002; Diester and Nieder, 2008; Hussar and Pasternak, 2012), visual attention in V4 (Mitchell et al., 2007), visual perceptual decision-making in the frontal eye field (Ding and Gold, 2011), motor control in the motor and premotor cortices (Isomura et al., 2009; Kaufman et al., 2010), and auditory processing during the passive listening in the auditory cortex (Atencio and Schreiner, 2008; Sakata and Harris, 2009; Ogawa et al., 2011). Interestingly, most of these studies showed differential roles in pyramidal neurons and interneurons.

Recently, using differences in the waveform of extracellularly-recorded neurons, we found that putative pyramidal neurons and interneurons in the lateral belt differentially encode and represent auditory categories (Tsunada et al., 2012). Specifically, we found that interneurons, on average, are more sensitive for auditory-category information than pyramidal neurons, although both neuron classes reliably encode category information (Figure 4).

Unfortunately, to our knowledge, there have not been other auditory-category studies that have examined the relative category sensitivity of pyramidal neurons vs. interneurons. However, a comparable visual-categorization study on numerosity in the

prefrontal cortex (Diester and Nieder, 2008) provides an opportunity to compare results across studies. Unlike our finding, Diester and Nieder found greater category sensitivity for putative pyramidal neurons than for putative interneurons.

The bases for these different sets of findings are unclear. However, three non-exclusive possibilities may underlie these differences. One possibility may relate to differences in the local-connectivity patterns and interactions between pyramidal neurons and interneurons across cortical areas (Wilson et al., 1994; Constantinidis and Goldman-Rakic, 2002; Diester and Nieder, 2008; Kätzel et al., 2010; Tsunada et al., 2012). Indeed, in the prefrontal cortex, simultaneously recorded (and, hence, nearby)



pyramidal neurons and interneurons have different category preferences (Diester and Nieder, 2008). In contrast, in the auditory cortex, simultaneously recorded pairs of pyramidal neurons and interneurons have similar category preferences (Tsunada et al., 2012). Thus, there may be different mechanisms for shaping category sensitivity across cortical areas. Second, the nature of the categorization task may also affect, in part, the category sensitivity of pyramidal neurons and interneurons: our task was a relatively simple task requiring the categorization of speech sounds based primarily on perceptual similarity, whereas Diester and Nieder's study required a more abstract categorization of numerosity. Finally, the third possibility relates to differences between stimulus dynamics: the visual stimuli in the Diester and Nieder's study were static stimuli, whereas our speech sounds had a rich spectrotemporal dynamic structure. To categorize dynamic stimuli, the moment-by-moment features of stimuli need to be quickly categorized. Thus, the greater category sensitivity of interneurons along with their well-known inhibitory influence on pyramidal neurons (Hefti and Smith, 2003; Wehr and Zador, 2003; Atencio and Schreiner, 2008; Fino and Yuste, 2011; Isaacson and Scanziani, 2011; Packer and Yuste, 2011; Zhang et al., 2011) may underlie the neural computations needed to create categorical representations of dynamic stimuli in the auditory cortex.

CONCLUSIONS AND FUTURE DIRECTIONS

Different neural transformations across different scales of neural organization progress during auditory categorization. Along the ascending auditory system in the ventral pathway, there is a progression in the encoding of categories from simple acoustic categories to categories representing abstract information. On the other hand, in local microcircuits within a cortical area, different classes of neurons, pyramidal neurons and interneurons, differentially compute categorical information. The computation is likely dependent upon the functional organization of the cortical area and dynamics of stimuli.

Despite several advances in our understanding of neural mechanism of auditory categorization, there still remain many important questions to be addressed. For example, it is poorly understood how bottom-up inputs from hierarchically lower areas, top-down feedback from higher areas, and local computations interact to form neural representations of auditory categories. Answering this question will provide a more thorough understanding of the information flow in the ventral auditory pathway. Another important question to be tested is what neural circuit mechanisms produce different category sensitivity between pyramidal neurons and interneurons, and functional roles of pyramidal neurons and interneurons in auditory categorization. Relevant to this question, the role that cortical laminae (another key element of local microcircuitry) play in auditory categorization should be also tested. Recent advances in experimental and analysis techniques should enable us to clarify the functional role of different classes of neurons in auditory categorization (Letzkus et al., 2011; Znamenskiy and Zador, 2013) and also test neural categorization across cortical layers (Lakatos et al., 2008; Takeuchi et al., 2011), providing further insights for neural computations for auditory categorization within local microcircuits.

ACKNOWLEDGMENTS

We thank Kate Christison-Lagay, Steven Eliades, and Heather Hersh for helpful comments on the preparation of this manuscript. We also thank Brian Russ and Jung Lee for data collection and Harry Shirley for outstanding veterinary support in our previous experiments. Joji Tsunada and Yale E. Cohen were supported by grants from NIDCD-NIH and the Boucai Hearing Restoration Fund.

REFERENCES

- Altmann, C. F., Doehrmann, O., and Kaiser, J. (2007). Selectivity for animal vocalizations in the human auditory cortex. *Cereb. Cortex* 17, 2601–2608. doi: 10.1093/cercor/bhl167
- Atencio, C. A., and Schreiner, C. E. (2008). Spectrotemporal processing differences between auditory cortical fast-spiking and regular-spiking neurons. *J. Neurosci.* 28, 3897–3910. doi: 10.1523/JNEUROSCI.5366-07.2008
- Averbeck, B. B., and Romanski, L. M. (2004). Principal and independent components of macaque vocalizations: constructing stimuli to probe high-level sensory processing. *J. Neurophysiol.* 91, 2897–2909. doi: 10.1152/jn.01103.2003
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–311. doi: 10.1038/35002078
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528. doi: 10.1093/cercor/10.5.512
- Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., and Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nat. Neurosci.* 7, 295–301. doi: 10.1038/nn1198
- Bizley, J. K., and Cohen, Y. E. (2013). The what, where, and how of auditory-object perception. *Nat. Rev. Neurosci.* 14, 693–707. doi: 10.1038/nrn3565
- Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., and Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nat. Neurosci.* 13, 1428–1432. doi: 10.1038/nn.2641
- Cohen, Y. E., Hauser, M. D., and Russ, B. E. (2006). Spontaneous processing of abstract categorical information in the ventrolateral prefrontal cortex. *Biol. Lett.* 2, 261–265. doi: 10.1098/rsbl.2005.0436
- Cohen, Y. E., Russ, B. E., Davis, S. J., Baker, A. E., Ackelson, A. L., and Nitecki, R. (2009). A functional role for the ventrolateral prefrontal cortex in non-spatial auditory cognition. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20045–20050. doi: 10.1073/pnas.0907248106
- Cohen, Y. E., Theunissen, F., Russ, B. E., and Gill, P. (2007). Acoustic features of rhesus vocalizations and their representation in the ventrolateral prefrontal cortex. *J. Neurophysiol.* 97, 1470–1484. doi: 10.1152/jn.00769.2006
- Constantinidis, C., and Goldman-Rakic, P. S. (2002). Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. *J. Neurophysiol.* 88, 3487–3497. doi: 10.1152/jn.00188.2002
- Desai, R., Liebenthal, E., Waldron, E., and Binder, J. R. (2008). Left posterior temporal regions are sensitive to auditory categorization. *J. Cogn. Neurosci.* 20, 1174–1188. doi: 10.1162/jocn.2008.20081
- Dewitt, I., and Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proc. Natl. Acad. Sci. U.S.A.* 109, E505–E514. doi: 10.1073/pnas.1113427109
- Dicarlo, J. J., Zoccolan, D., and Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron* 73, 415–434. doi: 10.1016/j.neuron.2012.01.010
- Diester, I., and Nieder, A. (2008). Complementary contributions of prefrontal neuron classes in abstract numerical categorization. *J. Neurosci.* 28, 7737–7747. doi: 10.1523/JNEUROSCI.1347-08.2008
- Ding, L., and Gold, J. I. (2011). Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. *Cereb. Cortex* 22, 1052–1067. doi: 10.1093/cercor/bhr178
- Doehrmann, O., Naumer, M. J., Volz, S., Kaiser, J., and Altmann, C. F. (2008). Probing category selectivity for environmental sounds in the human auditory brain. *Neuropsychologia* 46, 2776–2786. doi: 10.1016/j.neuropsychologia.2008.05.011
- Eckardt, W., and Zuberbühler, K. (2004). Cooperation and competition in two forest monkeys. *Behav. Ecol.* 15, 400–411. doi: 10.1093/beheco/arh032

- Engel, L. R., Frum, C., Puce, A., Walker, N. A., and Lewis, J. W. (2009). Different categories of living and non-living sound-sources activate distinct cortical networks. *Neuroimage* 47, 1778–1791. doi: 10.1016/j.neuroimage.2009.05.041
- Engineer, C. T., Perez, C. A., Carraway, R. S., Chang, K. Q., Roland, J. L., Sloan, A. M., et al. (2013). Similarity of cortical activity patterns predicts generalization behavior. *PLoS ONE* 8:e78607. doi: 10.1371/journal.pone.0078607
- Engineer, C. T., Perez, C. A., Chen, Y. H., Carraway, R. S., Reed, A. C., Shetake, J. A., et al. (2008). Cortical activity patterns predict speech discrimination ability. *Nat. Neurosci.* 11, 603–608. doi: 10.1038/nn.2109
- Fecteau, S., Armony, J. L., Joannette, Y., and Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *Neuroimage* 23, 840–848. doi: 10.1016/j.neuroimage.2004.09.019
- Fecteau, S., Armony, J. L., Joannette, Y., and Belin, P. (2005). Sensitivity to voice in human prefrontal cortex. *J. Neurophysiol.* 94, 2251–2254. doi: 10.1152/jn.00329.2005
- Fino, E., and Yuste, R. (2011). Dense inhibitory connectivity in neocortex. *Neuron* 69, 1188–1203. doi: 10.1016/j.neuron.2011.02.025
- Freedman, D. J., and Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature* 443, 85–88. doi: 10.1038/nature05078
- Freedman, D. J., and Miller, E. K. (2008). Neural mechanisms of visual categorization: insights from neurophysiology. *Neurosci. Biobehav. Rev.* 32, 311–329. doi: 10.1016/j.neubiorev.2007.07.011
- Freedman, D. J., Riesenhuber, M., Poggio, T., and Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291, 312–316. doi: 10.1126/science.291.5502.312
- Gifford, G. W. 3rd., Maclean, K. A., Hauser, M. D., and Cohen, Y. E. (2005). The neurophysiology of functionally meaningful categories: macaque ventrolateral prefrontal cortex plays a critical role in spontaneous categorization of species-specific vocalizations. *J. Cogn. Neurosci.* 17, 1471–1482. doi: 10.1162/0898929054985464
- Gifford, G. W. 3rd., Hauser, M. D., and Cohen, Y. E. (2003). Discrimination of functionally referential calls by laboratory-housed rhesus macaques: implications for neuroethological studies. *Brain Behav. Evol.* 61, 213–224. doi: 10.1159/000070704
- Giordano, B. L., McAdams, S., Kriegeskorte, N., Zatorre, R. J., and Belin, P. (2013). Abstract encoding of auditory objects in cortical activity patterns. *Cereb. Cortex* 23, 2025–2037. doi: 10.1093/cercor/bhs162
- Gold, J. I., and Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574. doi: 10.1146/annurev.neuro.29.051605.113038
- Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron* 14, 477–485. doi: 10.1016/0896-6273(95)90304-6
- Goll, J. C., Crutch, S. J., and Warren, J. D. (2010). Central auditory disorders: toward a neuropsychology of auditory objects. *Curr. Opin. Neurol.* 23, 617–627. doi: 10.1097/WCO.0b013e32834027f6
- González-Burgos, G., Krimer, L. S., Povyshva, N. V., Barrionuevo, G., and Lewis, D. A. (2005). Functional properties of fast spiking interneurons and their synaptic connections with pyramidal cells in primate dorsolateral prefrontal cortex. *J. Neurophysiol.* 93, 942–953. doi: 10.1152/jn.00787.2004
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: John Wiley and Sons, Inc.
- Hackett, T. A., Stepniowska, L., and Kaas, J. H. (1998). Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *J. Comp. Neurol.* 394, 475–495. doi: 10.1002/(SICI)1096-9861(19980518)394:4<475::AID-CNE6>3.0.CO;2-Z
- Hauser, M. D. (1998). Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Anim. Behav.* 55, 1647–1658. doi: 10.1006/anbe.1997.0712
- Hauser, M. D., and Marler, P. (1993a). Food-associated calls in rhesus macaques (*Macaca mulatta*) I. Socioecological factors influencing call production. *Behav. Ecol.* 4, 194–205. doi: 10.1093/beheco/4.3.194
- Hauser, M. D., and Marler, P. (1993b). Food-associated calls in rhesus macaques (*Macaca mulatta*) II. Costs and benefits of call production and suppression. *Behav. Ecol.* 4, 206–212. doi: 10.1093/beheco/4.3.206
- Hefti, B. J., and Smith, P. H. (2003). Distribution and kinetic properties of GABAergic inputs to layer V pyramidal cells in rat auditory cortex. *J. Assoc. Res. Otolaryngol.* 4, 106–121. doi: 10.1007/s10162-002-3012-z
- Holt, L. L., and Lotto, A. J. (2010). Speech perception as categorization. *Atten. Percept. Psychophys.* 72, 1218–1227. doi: 10.3758/APP.72.5.1218
- Hussar, C. R., and Pasternak, T. (2012). Memory-guided sensory comparisons in the prefrontal cortex: contribution of putative pyramidal cells and interneurons. *J. Neurosci.* 32, 2747–2761. doi: 10.1523/JNEUROSCI.5135-11.2012
- Hu, Z., Zhang, R., Zhang, Q., Liu, Q., and Li, H. (2012). Neural correlates of audio-visual integration of semantic category information. *Brain Lang.* 121, 70–75. doi: 10.1016/j.bandl.2012.01.002
- Isaacson, J. S., and Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron* 72, 231–243. doi: 10.1016/j.neuron.2011.09.027
- Isomura, Y., Harukuni, R., Takekawa, T., Aizawa, H., and Fukui, T. (2009). Microcircuitry coordination of cortical motor information in self-initiation of voluntary movements. *Nat. Neurosci.* 12, 1586–1593. doi: 10.1038/nn.2431
- Joassin, F., Maurage, P., and Campanella, S. (2011). The neural network sustaining the crossmodal processing of human gender from faces and voices: an fMRI study. *Neuroimage* 54, 1654–1661. doi: 10.1016/j.neuroimage.2010.08.073
- Kaas, J. H., and Hackett, T. A. (1999). “What” and “where” processing in auditory cortex. *Nat. Neurosci.* 2, 1045–1047. doi: 10.1038/15967
- Kaas, J. H., and Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11793–11799. doi: 10.1073/pnas.97.22.11793
- Kaas, J. H., Hackett, T. A., and Tramo, M. J. (1999). Auditory processing in primate cerebral cortex. *Curr. Opin. Neurobiol.* 9, 164–170. doi: 10.1016/S0959-4388(99)80022-1
- Kätzel, D., Zemelmann, B. V., Buetfering, C., Wölfel, M., and Miesenböck, G. (2010). The columnar and laminar organization of inhibitory connections to neocortical excitatory cells. *Nat. Neurosci.* 14, 100–107. doi: 10.1038/nn.2687
- Kaufman, M. T., Churchland, M. M., Santhanam, G., Yu, B. M., Afshar, A., Ryu, S. L., et al. (2010). Roles of monkey premotor neuron classes in movement preparation and execution. *J. Neurophysiol.* 104, 799–810. doi: 10.1152/jn.00231.2009
- Kawaguchi, Y., and Kondo, S. (2002). Parvalbumin, somatostatin and cholecystokinin as chemical markers for specific GABAergic interneuron types in the rat frontal cortex. *J. Neurocytol.* 31, 277–287. doi: 10.1023/A:1024126110356
- Kawaguchi, Y., and Kubota, Y. (1993). Correlation of physiological subgroupings of nonpyramidal cells with parvalbumin- and calbindinD28k-immunoreactive neurons in layer V of rat frontal cortex. *J. Neurophysiol.* 70, 387–396.
- Kawaguchi, Y., and Kubota, Y. (1997). GABAergic cell subtypes and their synaptic connections in rat frontal cortex. *Cereb. Cortex* 7, 476–486. doi: 10.1093/cercor/7.6.476
- Kluender, K. R., Diehl, R. L., and Killeen, P. (1987). Japanese quail can learn phonetic categories. *Science* 237, 1195–1197. doi: 10.1126/science.3629235
- Kuhl, P. K., and Miller, J. D. (1975). Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science* 190, 69–72. doi: 10.1126/science.1166301
- Kuhl, P. K., and Miller, J. D. (1978). Speech perception by the chinchilla: identification function for synthetic VOT stimuli. *J. Acoust. Soc. Am.* 63, 905–917. doi: 10.1121/1.381770
- Kuhl, P. K., and Padden, D. M. (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Percept. Psychophys.* 32, 542–550. doi: 10.3758/BF03204208
- Kuhl, P. K., and Padden, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *J. Acoust. Soc. Am.* 73, 1003–1010. doi: 10.1121/1.389148
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113. doi: 10.1126/science.1154735
- Leaver, A. M., and Rauschecker, J. P. (2010). Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *J. Neurosci.* 30, 7604–7612. doi: 10.1523/JNEUROSCI.0296-10.2010
- Letzkus, J. J., Wolff, S. B., Meyer, E. M., Tovote, P., Courtin, J., Herry, C., et al. (2011). A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* 480, 331–335. doi: 10.1038/nature10674
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol. Rev.* 5, 552–563. doi: 10.1037/h0020279
- Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., and Medler, D. A. (2005). Neural substrates of phonemic perception. *Cereb. Cortex* 15, 1621–1631. doi: 10.1093/cercor/bhi040

- Liebenthal, E., Desai, R., Ellingson, M. M., Ramachandran, B., Desai, A., and Binder, J. R. (2010). Specialization along the left superior temporal sulcus for auditory categorization. *Cereb. Cortex* 20, 2958–2970. doi: 10.1093/cercor/bhq045
- Logothetis, N. K., and Sheinberg, D. L. (1996). Visual object recognition. *Annu. Rev. Neurosci.* 19, 577–621. doi: 10.1146/annurev.ne.19.030196.003045
- Lotto, A. J., Kluender, K. R., and Holt, L. L. (1997). Perceptual compensation for coarticulation by Japanese quail (*Coturnix coturnix japonica*). *J. Acoust. Soc. Am.* 102, 1134–1140. doi: 10.1121/1.419865
- Mann, V. A. (1980). Influence of preceding liquid on stop-consonant perception. *Percept. Psychophys.* 28, 407–412. doi: 10.3758/BF03204884
- Markham, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneuron of the neocortical inhibitory system. *Nat. Rev. Neurosci.* 5, 793–807. doi: 10.1038/nrn1519
- McCormick, D. A., Connors, B. W., Lighthall, J. W., and Prince, D. A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *J. Neurophysiol.* 54, 782–806.
- Merchant, H., De Lafuente, V., Pena-Ortega, F., and Larriva-Sahd, J. (2012). Functional impact of interneuronal inhibition in the cerebral cortex of behaving animals. *Prog. Neurobiol.* 99, 163–178. doi: 10.1016/j.pneurobio.2012.08.005
- Mesgarani, N., Cheung, C., Johnson, K., and Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science* 343, 1006–1010. doi: 10.1126/science.1245994
- Mesgarani, N., David, S. V., Fritz, J. B., and Shamma, S. A. (2008). Phoneme representation and classification in primary auditory cortex. *J. Acoust. Soc. Am.* 123, 899–909. doi: 10.1121/1.2816572
- Miller, C. T., and Cohen, Y. E. (2010). “Vocalization processing,” in *Primate Neuroethology*, eds A. Ghazanfar and M. L. Platt (Oxford, UK: Oxford University Press), 237–255. doi: 10.1093/acprof:oso/9780195326598.003.0013
- Miller, E., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167
- Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nat. Rev. Neurosci.* 1, 59–65. doi: 10.1038/35036228
- Miller, E. K., Freedman, D. J., and Wallis, J. D. (2002). The prefrontal cortex: categories, concepts, and cognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 29, 1123–1136. doi: 10.1098/rstb.2002.1099
- Miller, E. K., Nieder, A., Freedman, D. J., and Wallis, J. D. (2003). Neural correlates of categories and concepts. *Curr. Opin. Neurobiol.* 13, 198–203. doi: 10.1016/S0959-4388(03)00037-0
- Mitchell, J. F., Sundberg, K. A., and Reynolds, J. H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* 5, 131–141. doi: 10.1016/j.neuron.2007.06.018
- Mottron, R., Calvert, G. A., Jaaskelainen, I. P., Matthews, P. M., Thesen, T., Tuomainen, J., et al. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage* 30, 563–569. doi: 10.1016/j.neuroimage.2005.10.002
- Mountcastle, V. B., Talbot, W. H., Sakata, H., and Hyvarinen, J. (1969). Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys. Neuronal periodicity and frequency discrimination. *J. Neurophysiol.* 32, 452–484.
- Nieder, A. (2012). Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11860–11865. doi: 10.1073/pnas.1204580109
- Nourski, K. V., Reale, R. A., Oya, H., Kawasaki, H., Kovach, C. K., Chen, H., et al. (2009). Temporal envelope of time-compressed speech represented in the human auditory cortex. *J. Neurosci.* 29, 15564–15574. doi: 10.1523/JNEUROSCI.3065-09.2009
- Obleser, J., Boecker, H., Drzezga, A., Haslinger, B., Hennenlotter, A., Roetinger, M., et al. (2006). Vowel sound extraction in anterior superior temporal cortex. *Hum. Brain Mapp.* 27, 562–571. doi: 10.1002/hbm.20201
- Obleser, J., Leaver, A. M., Van Meter, J., and Rauschecker, J. P. (2010). Segregation of vowels and consonants in human auditory cortex: evidence for distributed hierarchical organization. *Front. Psychol.* 1:232. doi: 10.3389/fpsyg.2010.00232
- Obleser, J., Zimmermann, J., Van Meter, J., and Rauschecker, J. P. (2007). Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb. Cortex* 17, 2251–2257. doi: 10.1093/cercor/bhl133
- Ogawa, T., Riera, J., Goto, T., Sumiyoshi, A., Noaka, H., Jerbi, K., et al. (2011). Large-scale heterogeneous representation of sound attributes in rat primary auditory cortex: from unit activity to population dynamics. *J. Neurosci.* 31, 14639–14653. doi: 10.1523/JNEUROSCI.0086-11.2011
- Ohl, F. W., Scheich, H., and Freeman, W. J. (2001). Change in pattern of ongoing cortical activity with auditory category learning. *Nature* 412, 733–736. doi: 10.1038/35089076
- Osada, T., Adachi, Y., Kimura, H. M., and Miyashita, Y. (2008). Towards understanding of the cortical network underlying associative memory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 2187–2199. doi: 10.1098/rstb.2008.2271
- Packer, A. M., and Yuste, R. (2011). Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *J. Neurosci.* 31, 13260–13271. doi: 10.1523/JNEUROSCI.3131-11.2011
- Pastore, R. E., Li, X. F., and Layer, J. K. (1990). Categorical perception of nonspeech chirps and bleats. *Percept. Psychophys.* 48, 151–156. doi: 10.3758/BF03207082
- Perrodin, C., Kayser, C., Logothetis, N. K., and Petkov, C. I. (2011). Voice cells in the primate temporal lobe. *Curr. Biol.* 21, 1408–1415. doi: 10.1016/j.cub.2011.07.028
- Petkov, C. I., Kayser, C., Studel, T., Whittingstall, K., Augath, M., and Logothetis, N. K. (2008). A voice region in the monkey brain. *Nat. Neurosci.* 11, 367–374. doi: 10.1038/nn2043
- Plakke, B., Diltz, M. D., and Romanski, L. M. (2013a). Coding of vocalizations by single neurons in ventrolateral prefrontal cortex. *Hear. Res.* 305, 135–143. doi: 10.1016/j.heares.2013.07.011
- Plakke, B., Hwang, J., Diltz, M. D., and Romanski, L. M. (2013b). “The role of ventral prefrontal cortex in auditory, visual and audiovisual working memory,” in *Society for Neuroscience, Program No. 574.515 Neuroscience Meeting Planner* (San Diego, CA).
- Plakke, B., Ng, C. W., and Poremba, A. (2013c). Neural correlates of auditory recognition memory in primate lateral prefrontal cortex. *Neuroscience* 244, 62–76. doi: 10.1016/j.neuroscience.2013.04.002
- Poremba, A., Bigelow, J., and Rossi, B. (2013). Processing of communication sounds: contributions of learning, memory, and experience. *Hear. Res.* 305, 31–44. doi: 10.1016/j.heares.2013.06.005
- Rao, S. G., Williams, G. V., and Goldman-Rakic, P. S. (1999). Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in PFC. *J. Neurophysiol.* 81, 1903–1915.
- Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Curr. Opin. Neurobiol.* 8, 516–521. doi: 10.1016/S0959-4388(98)80040-8
- Rauschecker, J. P. (2012). Ventral and dorsal streams in the evolution of speech and language. *Front. Evol. Neurosci.* 4:7. doi: 10.3389/fnevo.2012.00007
- Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724. doi: 10.1038/nn.2331
- Rauschecker, J. P., and Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11800–11806. doi: 10.1073/pnas.97.22.11800
- Rauschecker, J. P., Tian, B., and Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268, 111–114. doi: 10.1126/science.7701330
- Recanzone, G. H., and Cohen, Y. E. (2010). Serial and parallel processing in the primate auditory cortex revisited. *Behav. Brain Res.* 5, 1–6. doi: 10.1016/j.bbr.2009.08.015
- Romanski, L. M., and Averbeck, B. B. (2009). The primate cortical auditory system and neural representation of conspecific vocalizations. *Annu. Rev. Neurosci.* 32, 315–346. doi: 10.1146/annurev.neuro.051508.135431
- Romanski, L. M., Bates, J. F., and Goldman-Rakic, P. S. (1999a). Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* 403, 141–157. doi: 10.1002/(SICI)1096-9861(19990111)403:2<141::AID-CNE1>3.0.CO;2-V
- Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., and Rauschecker, J. P. (1999b). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136. doi: 10.1038/16056
- Russ, B. E., Ackelson, A. L., Baker, A. E., and Cohen, Y. E. (2008a). Coding of auditory-stimulus identity in the auditory non-spatial processing stream. *J. Neurophysiol.* 99, 87–95. doi: 10.1152/jn.01069.2007
- Russ, B. E., Lee, Y.-S., and Cohen, Y. E. (2007). Neural and behavioral correlates of auditory categorization. *Hear. Res.* 229, 204–212. doi: 10.1016/j.heares.2006.10.010

- Russ, B. E., Orr, L. E., and Cohen, Y. E. (2008b). Prefrontal neurons predict choices during an auditory same-different task. *Curr. Biol.* 18, 1483–1488. doi: 10.1016/j.cub.2008.08.054
- Sakata, S., and Harris, K. D. (2009). Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex. *Neuron* 64, 404–418. doi: 10.1016/j.neuron.2009.09.020
- Schreiner, C. E. (1998). Spatial distribution of responses to simple and complex sounds in the primary auditory cortex. *Audiol. Neurootol.* 3, 104–122. doi: 10.1159/000013785
- Seleznova, E., Scheich, H., and Brosch, M. (2006). Dual time scales for categorical decision making in auditory cortex. *Curr. Biol.* 16, 2428–2433. doi: 10.1016/j.cub.2006.10.027
- Sinnott, J. M., and Brown, C. H. (1997). Perception of the American English liquid vertical bar ra-la vertical bar contrast by humans and monkeys. *J. Acoust. Soc. Am.* 102, 588–602. doi: 10.1121/1.419732
- Staeren, N., Renvall, H., De Martino, F., Goebel, R., and Formisano, E. (2009). Sound categories are represented as distributed patterns in the human auditory cortex. *Curr. Biol.* 19, 498–502. doi: 10.1016/j.cub.2009.01.066
- Steinschneider, M. (2013). “Phonemic representations and categories,” in *Neural Correlates of Auditory Cognition*, eds Y. E. Cohn, A. N. Popper and R. R. Fay (New York, NY: Springer), 151–191. doi: 10.1007/978-1-4614-2350-8_6
- Steinschneider, M., Fishman, Y. I., and Arezzo, J. C. (2003). Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *J. Acoust. Soc. Am.* 114, 307–321. doi: 10.1121/1.1582449
- Steinschneider, M., Nourski, K. V., Kawasaki, H., Oya, H., Brugge, J. F., and Howard, M. A. (2011). Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. *Cereb. Cortex* 10, 2332–2347. doi: 10.1093/cercor/bhr014
- Steinschneider, M., Volkov, I. O., Fishman, Y. I., Oya, H., Arezzo, J. C., and Howard, M. A. 3rd. (2005). Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. *Cereb. Cortex* 15, 170–186. doi: 10.1093/cercor/bhh120
- Takeuchi, D., Hirabayashi, T., Tamura, K., and Miyashita, Y. (2011). Reversal of interlaminar signal between sensory and memory processing in monkey temporal cortex. *Science* 331, 1443–1447. doi: 10.1126/science.1199967
- Tian, B., and Rauschecker, J. P. (2004). Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. *J. Neurophysiol.* 92, 2993–3013. doi: 10.1152/jn.00472.2003
- Tsunada, J., Lee, J. H., and Cohen, Y. E. (2011). Representation of speech categories in the primate auditory cortex. *J. Neurophysiol.* 105, 2634–2646. doi: 10.1152/jn.00037.2011
- Tsunada, J., Lee, J. H., and Cohen, Y. E. (2012). Differential representation of auditory categories between cell classes in primate auditory cortex. *J. Physiol.* 590, 3129–3139. doi: 10.1113/jphysiol.2012.232892
- Van Lancker, D. R., and Canter, G. J. (1982). Impairment of voice and face recognition in patients with hemispheric damage. *Brain Cogn.* 1, 185–195.
- Van Lancker, D. R., Cummings, J. L., Kreiman, J., and Dobkin, B. H. (1988). Phonagnosia—a dissociation between familiar and unfamiliar voices. *Cortex* 24, 195–209. doi: 10.1016/S0010-9452(88)80029-7
- Warren, J. D., Scott, S. K., Price, C. J., and Griffiths, T. D. (2006). Human brain mechanisms for the early analysis of voices. *Neuroimage* 31, 1389–1397. doi: 10.1016/j.neuroimage.2006.01.034
- Wehr, M. S., and Zador, A. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 27, 442–446. doi: 10.1038/nature02116
- Werner, S., and Noppeney, U. (2010). Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *J. Neurosci.* 30, 2662–2675. doi: 10.1523/JNEUROSCI.5091-09.2010
- Wilson, F. A., O’Scalaidhe, S. P., and Goldman-Rakic, P. S. (1994). Functional synergism between putative gamma-aminobutyrate-containing neurons and pyramidal neurons in prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 91, 4009–4013. doi: 10.1073/pnas.91.9.4009
- Zhang, L. I., Zhou, Y., and Tao, H. W. (2011). Perspectives on: information and coding in mammalian sensory physiology: inhibitory synaptic mechanisms underlying functional diversity in auditory cortex. *J. Gen. Physiol.* 138, 311–320. doi: 10.1085/jgp.201110650
- Znamenskiy, P., and Zador, A. M. (2013). Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* 497, 482–485. doi: 10.1038/nature12077
- Züberbühler, K. (2000a). Causal cognition in a non-human primate: field playback experiments with Diana monkeys. *Cognition* 76, 195–207. doi: 10.1016/S0010-0277(00)00079-2
- Züberbühler, K. (2000b). Interspecies semantic communication in two forest primates. *Proc. R. Soc. Lond. B Biol. Sci.* 267, 713–718. doi: 10.1098/rspb.2000.1061
- Zuberbühler, K., and Seyfarth, R. M. (1997). Diana monkey long-distance calls: messages for conspecifics and predators. *Anim. Behav.* 53, 589–604. doi: 10.1006/anbe.1996.0334

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 March 2014; accepted: 27 May 2014; published online: 17 June 2014.

Citation: Tsunada J and Cohen YE (2014) Neural mechanisms of auditory categorization: from across brain areas to within local microcircuits. *Front. Neurosci.* 8:161. doi: 10.3389/fnins.2014.00161

This article was submitted to *Auditory Cognitive Neuroscience*, a section of the journal *Frontiers in Neuroscience*.

Copyright © 2014 Tsunada and Cohen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.