

# Neurocomputational Mechanisms Contributing to Auditory Perception

Yale E. Cohen<sup>1,2,3</sup>, Taku Banno<sup>1</sup>, Jaejin Lee<sup>1</sup>, Francisco Rodriguez-Campos<sup>1</sup>, Matthew Schaff<sup>4</sup>, Lalitta Suriya-Arunroj<sup>1</sup>, Joji Tsunada<sup>1</sup>

<sup>1</sup> Department of Otorhinolaryngology, University of Pennsylvania, Philadelphia, PA 19104.  
ycohen@penncmedicine.upenn.edu

<sup>2</sup> Department of Neuroscience, University of Pennsylvania, Philadelphia, PA 19104

<sup>3</sup> Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104

<sup>4</sup> Neuroscience Graduate Group, University of Pennsylvania, Philadelphia, PA 19104

## Summary

A fundamental scientific question in auditory neuroscience is identifying the mechanisms required by the brain to transform an unlabelled mixture of auditory stimuli into distinct and coherent perceptual representations. This transformation is often called "auditory-scene analysis". Auditory-scene analysis consists of a complex interaction of multiple neurocomputational processes, including Gestalt grouping mechanisms, attention, and perceptual decision-making. Despite a great deal of scientific energy devoted to understanding these aspects of hearing, we still do not understand (1) how sound perception arises from neural activity and (2) the causal relationship between neural activity and sound perception. Several lines of evidence indicate that the ventral auditory pathway plays a prominent role in auditory perception and decision-making. Here, we review the contribution of the ventral pathway to auditory perception and put forth challenges to the field to further our understanding of the relationship between neural activity in the ventral pathway and perception.

© 2018 The Author(s). Published by S. Hirzel Verlag · EAA. This is an open access article under the terms of the Creative Commons Attribution (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/>).

PACS no. 43.64.-q, 43.64.Ri, 43.64.Tk

## 1. Introduction

Hearing is fundamental to and spans our human existence. We can hear selectively: we can focus on a sound that a banjo player is producing and ignore the music produced by the rest of the band. But, then, we can switch our attention to the voice of the band's singer. We can hear in a variety of environments: in a quiet library or in a busy, loud restaurant. Hearing in this loud restaurant is particularly challenging because our dining partner's voice is embedded in a complicated mixture of stimuli of other patrons talking, waiters describing the specials of the day, dishes clanking, and background music.

This remarkable capacity is associated with a number of computational processes, which act in both parallel and serial, including: perceptual grouping, decision-making, and attention. (1) Perceptual grouping is a form of feature-based stimulus segmentation that determines whether acoustic events are grouped into a single sound or are segregated into distinct sounds [1]. (2) Auditory decision-making is a computational process in which the

brain interprets sensory information in order to detect, discriminate, and identify the source or content of auditory stimuli [2]. (3) Although attention is not always necessary, our awareness of a sound can be influenced by attention [3]. For example, we can choose whether to listen to or ignore the mandolin, the fiddle, or even the whole band during a concert. Likewise, we can selectively attend to the particular features in a person's voice that allow us to identify the speaker.

It is widely believed that the neural computations and processes that mediate auditory perception are found in the ventral auditory pathway [4, 5, 6]. In rhesus monkeys [7, 8, 4], this pathway begins in the anterolateral (AL) belt region of auditory cortex, which receives input from core regions of auditory cortex, specifically the primary auditory cortex (A1) and field R and the mediolateral (ML) belt region. AL, in turn, projects directly and indirectly via the rostral parabelt (rPB) region of the auditory cortex, to the ventrolateral prefrontal cortex (vlPFC). An analogous pathway has been identified in humans [7].

In this review, we focus on the processing that occurs in different stages of the ventral auditory pathway. We review the mechanisms and representations along the ventral pathway, with an emphasis on those experiments that combined electrophysiology with auditory behavior. We inte-

---

Received 19 February 2018,  
accepted 27 August 2018.

grate these studies into a model of hierarchical processing in the ventral pathway. We attempt to reconcile different sets of studies and pose challenges to the field as to how to best study processing along this pathway.

## 2. Neural correlates of auditory perception in the auditory cortex

Although not technically a part of the ventral pathway, a large literature has focused on the contribution of A1 to perception and behavior. The results of these studies have been, to some degree, controversial. Typifying the controversy is A1's contribution to perceptual decision-making. In order to study decision-making, humans and non-human animals engage in a behavioral task, while neuronal signals are recorded simultaneously during the task. The critical manipulation is to probe behavior when a stimulus is "ambiguous". For example, if a listener is asked to report whether a sequence of tone bursts contained more "low-frequency" or "high-frequency" tone bursts (a "low-high" task), an ambiguous or "noisy" stimulus would be one that contained 50% low-frequency tone-bursts and 50% high-frequency tone bursts. Using this strategy, changes in stimulus features can be separated from those related to changes in the behavioral report. In essence, if a neuron codes a stimulus feature, its response should be invariant to the behavioral choice. In contrast, if a neuron reflects the behavioral choice, then neuronal activity should differ as a function of the choice. This is often called "choice-related" activity.

Several studies have noted such choice-related activity in A1 [9, 10, 11, 12, 13, 14]. For example, in one study in which ferrets participated in a pitch discrimination task, both local-field potentials and spiking activity were modulated by the ferrets' pitch judgments. Work from the Sutter lab has also identified A1 and belt-region activity that was modulated by choice; in these experiments, rhesus monkeys detected changes in amplitude-modulated noise. In a human neuroimaging study, core auditory cortex was modulated by whether listeners could discriminate between ambiguous speech sounds.

A different set of literature, however, has failed to identify choice-related activity in A1 and several belt regions of the auditory cortex. Work from Romo's group failed to see choice activity in A1 while rhesus monkeys were discriminating between acoustic flutter but did see elements of choice behavior in the ventral premotor cortex [15, 16]. Similarly, several studies from our lab have not identified choice-related activity in the auditory cortex [17, 18, 19]. In a recent study, we asked monkeys to detect a target stimulus in various signal-to-noise regimes and found that, on average, A1 activity was not modulated by choice. A different finding emerged at the population level: a linear decoder could read-out choice as well as various task parameters (single-to-noise level and "signal" versus "noise" trials) from these neurons. This implies that information about choice and task is available in the population, which

potentially could be encoded more robustly in those single neurons that receive input from A1. In another set of studies in which monkeys were engaged in the low-high task, we found average AL and ML activity was not modulated by choice. In a final study, monkeys participated in a two-interval, discrimination task in which they reported whether two phonemes were the "same" or "different". Once again, average AL activity was not modulated by choice.

We cannot totally reconcile these different sets of findings. However, we wonder whether these differences may relate to the specific nature of the different auditory decisions. For example, for those studies in which choice-related activity was observed [11, 12, 13], the animal listeners were asked to make an auditory decision about a relatively low-level stimulus feature (pitch or amplitude modulation). Because these features may be represented directly in an A1 neuron's firing rate, A1 activity may be able to encode the sensory evidence for these decisions. In contrast, in the sets of studies that did not identify choice-related activity, [15, 16, 17, 18, 19], rhesus monkeys were required to make a decision about a stimulus attribute that may not be encoded directly in a neuron's firing rate. Indeed, in our signal-to-noise detection task [19], we found that the firing rates of most A1 neurons were not modulated by different signal-to-noise levels. For such decisions, it is feasible that only later regions of the ventral auditory pathway participate in the decision process.

This raises an intriguing idea. Namely, the neurocomputations underlying auditory perception may be multiplexed along the ventral auditory pathway. In other words, depending on the demands of the task and the nature of the auditory decision, a particular brain region may differentially contribute to the different computations underlying auditory perception and decision-making. To test this idea, we would minimally have to record simultaneously in different brain regions and examine how single neurons and neuronal populations are modulated by the same task with different types of stimuli and also by different tasks with the same stimuli.

## 3. The contribution of time to understanding decision-making

We have previously reported the presence of choice-related activity in the vIPFC [20]. In that study, rhesus monkeys discriminated between two phonemes and reported whether they were the same or were different. We found that, for the same nominal stimulus, vIPFC activity was modulated by choice.

At first glance, this combination of studies from AL, ML, and vIPFC suggest a hierarchical processing stream in which sensory evidence early in the pathway (AL and ML) gets transformed into a categorical choice later in the pathway (vIPFC). However, although this posited hierarchy seems quite enticing, it is missing a critical piece of information: *when* do these computations occur? This timing information is critical because it can help us differenti-

ate between neuronal responses related to feedforward and feedback connectivity.

Indeed, the ventral auditory pathway is as rich in feedback connections to cortical and subcortical regions as it is in feedforward connections [7, 21]. Because of this rich interconnectivity, the presence of, for example, A1 choice-related activity does not necessarily imply that this information arises in A1 nor does it necessarily follow that it is part of a feedforward process in which this choice-related activity contributes causally to the eventual auditory decision. Similarly, if vIPFC choice-related activity occurs after the actual choice, our interpretation of its functional utility would be substantially different than if it occurred at the time of the choice.

One means to resolve this issue is to identify approaches by which a subject's behavioral responses can be related to underlying neurocomputational processes and ultimately to neuronal activity. A successful approach has been to combine behavioral tasks with variants of sequential-sampling models, like the drift-diffusion model [2]. These models quantify the process of converting incoming sensory evidence, which is represented in the brain as the noisy, spiking activity of populations of relevant sensory neurons, into a decision variable that guides behavior [22]. A key benefit of these models is that they can make quantitative predictions about both choice and response time as a function of the manipulated stimulus parameter (e.g., signal-to-noise ratio). Thus, we can use these models to jointly fit (1) the psychometric function, which describes accuracy versus the manipulated stimulus parameter and (2) the chronometric function, which describes response time versus the manipulated stimulus parameter. The fits to this model yield insights into the decision process that gives rise to the measured accuracy, response times, and trade-offs between the two [2]. Further, these model fits also allow us to approximate the time of occurrence of the decision process: the drift-diffusion model partitions a subject's response times into "decision" and "non-decision" times. Decision time reflects the time (relative to stimulus onset) needed to make a perceptual decision, whereas non-decision times reflect sensory delay, motor preparation and other post-decision processing.

Using this approach, we have made several important insights into the temporal dynamics of an auditory perceptual decision that are needed during the low-high task [17, 23]. These temporal dynamics reflect this particular auditory decision and whether or not they generalize to other auditory decisions is an open question. Regardless, relative to the time of stimulus onset, we found comparable stimulus sensitivity (i.e., the degree to which neuronal activity was modulated by the percentage of low- and high-frequency tone bursts in a sequence) and choice-related activity in ML and AL. However, on a neuron-by-neuron basis, when we compared AL's responses relative to the decision time, we found a correlation between stimulus sensitivity and choice-related activity. Importantly, this correlation became significant just *before* the decision time. We failed to find such a correlation in the ML data.

This suggests that AL—but not ML—may transmit the information used in this decision. Consistent with these observations, microstimulation of AL sites biased the rhesus monkeys' behavior, whereas ML sites did not bias behavior. vIPFC neurons appear, on average, to signal choice-related activity only *after* the inferred decision commitment. vIPFC may then play a role in monitoring or evaluating decision outcomes.

Where then is the brain region that actually encodes the decision? We have yet to identify this area. But, we suspect that the parabelt may be a prime target because it receives input from AL and projects to vIPFC. Even without the identification of the putative site(s) that encodes the actual decision, this temporal analysis strongly suggests a feedforward pathway between ML and the vIPFC along with the respective functional roles.

We believe that being able to identify the temporal dynamics of a decision-making process and then relating these dynamics to neuronal activity is a powerful way to identify a flow of information. We encourage others to use a similar approach in order to distinguish between feedforward and feedback contributions to perception and behavior.

#### 4. Conclusions and future questions

Of course, several fundamental questions remain. First, as noted above, understanding how feedforward versus feedback information contributes to neural correlates of perceptual judgments remains an open question. Relatedly, the degree to which subcortical processes interact with decision-making and these feedforward and feedback loops is an open question. Second, the degree to which different types of auditory judgements differentially engage different regions of the ventral pathway have yet to be fully articulated. A third question is to identify how the different computational processes (e.g., perceptual grouping, attention, and decision-making) that underlie auditory perception interact with one another at both the cortical and subcortical level. For example, it remains an open issue as to whether and how attention differentially modulates neural correlates of auditory perception at different hierarchical levels of the ventral auditory pathway (e.g., A1 versus AL) [24]. Another example is the potential interactions between auditory perceptual grouping and decision-making. Finally, it is important to identify the manner by which the dorsal and ventral auditory pathways interact in order to form a consistent and coherent representation of the auditory scene and the degree to which the dorsal pathway (and potential other pathways) mediate perception.

#### Acknowledgement

This work was supported by grants from the NIH, ARO, and ONR.

#### References

- [1] A. S. Bregman: Auditory scene analysis. MIT Press, Boston, MA, 1990.

- [2] J. I. Gold, M. N. Shadlen: The neural basis of decision making. *Annu Rev Neurosci* **30** (2007) 535–574.
- [3] B. G. Shinn-Cunningham: Object-based auditory and visual attention. *Trends Cogn Sci* **12** (2008) 182–186.
- [4] J. P. Rauschecker, B. Tian: Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc Natl Acad Sci U S A* **97** (2000) 11800–11806.
- [5] L. M. Romanski, B. B. Averbeck: The primate cortical auditory system and neural representation of conspecific vocalizations. *Annu Rev Neurosci* **32** (2009) 315–346.
- [6] J. K. Bizley, Y. Cohen: The what, where and how of auditory-object perception. *Nat Rev Neurosci* **14** (2013) 693–707.
- [7] T. A. Hackett: Information flow in the auditory cortical network. *Hear Res* **271** (2011) 133–146.
- [8] L. M. Romanski, B. Tian, J. Fritz, M. Mishkin, P. S. Goldman-Rakic, J. P. Rauschecker: Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci* **2** (1999) 1131–1136.
- [9] L. Riecke, D. Mendelsohn, C. Schreiner, E. Formisano: The continuity illusion adapts to the auditory scene. *Hear Res* **247** (2009) 71–77.
- [10] L. Riecke, M. Vanbussel, L. Hausfeld, D. Başkent, E. Formisano, F. Esposito: Hearing an illusory vowel in noise: suppression of auditory cortical activity. *J Neurosci* **32** (2012) 8024–8034.
- [11] M. Niwa, J. S. Johnson, K. N. O'Connor, M. L. Sutter: Activity related to perceptual judgment and action in primary auditory cortex. *J Neurosci* **32** (2012) 3193–3210.
- [12] M. Niwa, J. S. Johnson, K. N. O'Connor, M. L. Sutter: Differences between primary auditory cortex and auditory belt related to encoding and choice for am sounds. *J Neurosci* **33** (2013) 8378–8395.
- [13] J. K. Bizley, K. M. Walker, F. R. Nodal, A. J. King, J. W. Schnupp: Auditory cortex represents both pitch judgments and the corresponding acoustic cues. *Curr Biol* **23** (2013) 620–625.
- [14] N. Kilian-Hütten, G. Valente, J. Vroomen, E. Formisano: Auditory cortex encodes the perceptual interpretation of ambiguous sound. *J Neurosci* **31** (2011) 1715–1720.
- [15] L. Lemus, A. Hernandez, R. Romo: Neural codes for perceptual discrimination of acoustic flutter in the primate auditory cortex. *Proc Natl Acad Sci U S A* **106** (2009) 9471–9476.
- [16] L. Lemus, A. Hernandez, R. Romo: Neural encoding of auditory discrimination in ventral premotor cortex. *Proc Natl Acad Sci U S A* **106** (2009) 14640–14645.
- [17] J. Tsunada, A. S. K. Liu, J. I. Gold, Y. E. Cohen: Causal contribution of primate auditory cortex to auditory perceptual decision-making. *Nat Neurosci* **19** (2016) 135–142.
- [18] J. Tsunada, J. H. Lee, Y. Cohen: Representation of speech categories in the primate auditory cortex. *J Neurophysiol* **105** (2011) 2634–2646.
- [19] K. L. Christison-Lagay, S. Bennur, Y. E. Cohen: Contribution of spiking activity in the primary auditory cortex to detection in noise. *J Neurophysiol* **118** (2017) 3118–3131.
- [20] B. E. Russ, L. E. Orr, Y. E. Cohen: Prefrontal neurons predict choices during an auditory same-different task. *Curr Biol* **18** (2008) 1483–1488.
- [21] T. A. Hackett, L. A. de la Mothe, C. R. Camalier, A. Falchier, P. Lakatos, Y. Kajikawa, C. E. Schroeder: Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: refining the hierarchical model. *Front Neurosci* **8** (2014) 72.
- [22] A. J. Parker, W. T. Newsome: Sense and the single neuron: probing the physiology of perception. *Annual Review of Neuroscience* **21** (1998) 227–277.
- [23] J. Tsunada, A. S. K. Liu, J. I. Gold, Y. E. Cohen: Relative contributions of the auditory and prefrontal cortices to auditory perceptual decision-making. – In: *Neuroscience Meeting Planner*. Society for Neuroscience, Washington, D.C., 2014, 815.11.
- [24] E. selectivity for task-relevant stimuli in higher-order auditory cortex: Atiani, s. and david, s. v. and elgueda, d. and locastro, m. and radtke-schuller, s. and shamma, s. a. and fritz, j. b. *Neuron* **82** (2014) 486–499.