

Neural Mechanisms for the Abstraction and Use of Pitch Information in Auditory Cortex

Xiaoqin Wang^{1,2*} and Kerry M. M. Walker^{3*}

¹Tsinghua-Johns Hopkins Joint Center for Biomedical Engineering Research and Department of Biomedical Engineering, Tsinghua University, Beijing 100084, China, ²Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21025, and ³Department of Physiology, Anatomy & Genetics, University of Oxford, Oxford, OX1 3PT, United Kingdom

Experiments in animals have provided an important complement to human studies of pitch perception by revealing how the activity of individual neurons represents harmonic complex and periodic sounds. Such studies have shown that the acoustical parameters associated with pitch are represented by the spiking responses of neurons in A1 (primary auditory cortex) and various higher auditory cortical fields. The responses of these neurons are also modulated by the timbre of sounds. In marmosets, a distinct region on the low-frequency border of primary and non-primary auditory cortex may provide pitch tuning that generalizes across timbre classes.

Questions on pitch encoding mechanisms in auditory cortex

Animal studies have provided useful information at the neuronal level for understanding pitch processing mechanisms in auditory cortex. Auditory cortex in non-human primates contains a “core” region, composed of primary auditory cortex (A1), rostral area (R), and rostral-temporal area (RT) (Fig. 1*A,B*). The core is surrounded by “belt” and “para-belt” regions (Kaas and Hackett, 2000; Hackett et al., 2001). Other mammals also have a “core” region in auditory cortex [A1 and anterior auditory field area (AAF)], surrounded by secondary areas (ferret, Wallace et al., 1997; Bizley et al., 2005; cat, Winer and Lee, 2007) (Fig. 1*C,D*). The similarities in auditory cortex organization make it possible to compare data obtained from different species, although one also must bear in mind the differences among species during this practice.

There are several scenarios of how pitch could be encoded in auditory cortex. Cortical areas with an orderly tonotopic organization, like A1, can represent pitch in frequency regions corresponding to individual harmonic components, orthogonal to the tonotopic axis (Langer et al., 1997). Alternatively, pitch can be extracted by neurons tuned to low frequencies in a tonotopically organized cortical area (Bendor and Wang, 2005) or in a specialized cortical region (Griffiths and Hall, 2012). One would also like to know in any of such scenarios whether a representation of pitch can be generalized to all types of sounds bearing the same pitch (including missing fundamental sounds), and whether a particular neural representation is linked to pitch perception measured behaviorally. We will review below studies that address these questions.

Neurophysiological studies of pitch and periodicity coding in auditory cortex

Following the discovery that macaque monkeys (Tomlinson and Schwarz, 1988) and cats (Heffner and Whitfield, 1976) can perceive the pitch of missing fundamental harmonic complex sounds, neurophysiologists embarked on attempts to identify neurons in A1 that have the same periodicity tuning for pure tones and missing fundamental sounds. Initial attempts performed in awake monkeys were unsuccessful. Two research groups found that while the pure tone frequency tuning of many A1 neurons (i.e., characteristic frequency; CF) is consistent with the neuron's responses to harmonic tone complexes, their periodicity tuning does not extend to harmonic sounds with missing fundamentals (Schwarz and Tomlinson, 1990; Fishman et al., 1998). Such results dispelled early expectations of finding a single neuron explanation of pitch perception at the level of A1.

Several studies have, nevertheless, demonstrated that A1 neurons are tuned to a variety of acoustical cues that may be important prerequisites to pitch encoding. Neurons in A1 are arranged according to their CF, forming a tonotopic map of frequency preference across its surface. In the A1 of monkeys, the low-frequency harmonics of periodic click trains are represented as distinct areas of activation across the low-CF region of the tonotopic map (Steinschneider et al., 1998). This might serve to represent the resolved harmonics of periodic sounds (Oxenham, 2012). The temporal envelope modulations of higher (presumably unresolved) harmonics, on the other hand, are represented in the phase-locked responses of high-CF neurons (Steinschneider et al., 1998). Therefore, acoustical properties that are necessary for spectral- and temporal-based pitch extraction processes are represented across the population of neurons in A1.

A subpopulation of “multi-peaked” neurons have been identified in cats (Sutter and Schreiner, 1991) and marmosets (Kadia and Wang, 2003) which may be capable of harmonically fusing complex sounds. These neurons respond preferentially not just to a particular CF, but also to tones at frequencies that are harmonically related to their CF, particularly 1.5*CF and 2*CF. In

Received April 4, 2012; revised July 18, 2012; accepted July 23, 2012.

This work is supported by a grant from Tsinghua University (X.W.), NIH Grant R01 DC03180 (X.W.) and a Wellcome Trust Principal Research Fellowship (Andrew J. King).

*X.W. and K.M.M.W. contributed equally to this work.

Correspondence should be addressed to either of the following: Dr. Xiaoqin Wang, Department of Biomedical Engineering, Johns Hopkins University School of Medicine, 720 Rutland Avenue, Traylor 410, Baltimore, MD 21025, E-mail: xiaoqin.wang@jhu.edu, or Dr. Kerry Walker, Department of Physiology, Anatomy & Genetics, University of Oxford, Sherrington Building, Parks Road, Oxford, OX1 3PT, UK, E-mail: kerry.walker@dpag.ox.ac.uk.

DOI:10.1523/JNEUROSCI.3814-12.2012

Copyright © 2012 the authors 0270-6474/12/3213339-04\$15.00/0

some cases, spiking responses to the CF in the presence of its harmonics were enhanced over the response to CF alone (Kadia and Wang, 2003). Similarly, a small proportion (12%) of neurons in the two core auditory cortical fields of awake ferrets (A1 and AAF) have been shown to be harmonically sensitive (Kalluri et al., 2008). In these neurons, linear spectro-temporal filters computed from the cell's response to inharmonic sounds could not accurately predict the response to harmonic complex sounds. These studies show that multiple harmonic components of complex periodic sounds are integrated and represented as spike rate codes in a subpopulation of neurons in A1.

Extracellular recordings and intrinsic optical imaging of responses in gerbil and cat A1 to pure tones and sinusoidally amplitude-modulated (SAM) tones have suggested that a topographic gradient of best modulation frequencies may exist for complex, periodic sounds that is distinct from, or even orthogonal to, the tonotopic map (Schulze and Langner, 1997; Schulze et al., 2002; Langner et al., 2009). Topographic organization of best modulation rates has not yet been found in primate A1 (Schwarz and Tomlinson, 1990; Fishman et al., 1998). In an intrinsic optical imaging study of primary and secondary auditory cortices in ferrets, Nelken et al. (2008) did find a gradient of best modulation tuning for SAM tones, but here the gradient ran approximately parallel to the tonotopic map, in contrast to earlier studies in humans (Langner et al., 1997) and gerbils (Schulze et al., 2002). Moreover, although periodicity gradients were also observed for high-pass click trains and high-pass iterated rippled noise, there was no consistent arrangement of periodicity preference gradients across the three stimulus types (Nelken et al., 2008). The stimulus specificity of these periodicity preferences means that they cannot be interpreted as generalized pitch maps. Neuronal sensitivity to confounded features, such as harmonic spacing due to critical bandwidths (Fishman et al., 2000) or cochlear distortion products (Wiegand and Patterson, 1999), have been offered as potential explanations for the observed organization of modulation preferences for SAM tones.

Although A1 neurons can extract acoustical cues that are necessary to compute the pitch of a variety of periodic sounds, the response properties of a single A1 neuron would be insufficient to represent the pitch of the entire range of pitch-evoking sounds, particularly those with missing fundamentals. For most sounds that humans and animals encounter in their natural environments, such as vocal calls, these neural representations may be sufficient for pitch extraction. This possibility was examined by Bizley et al. (2010), who trained statistical "neurometric" algorithms to discriminate the periodicity of artificial vowel sounds (i.e., bandpass-filtered click trains) based on the responses of neurons in ferret auditory cortex. They found that neurometrics based on the response of small populations of auditory cortical

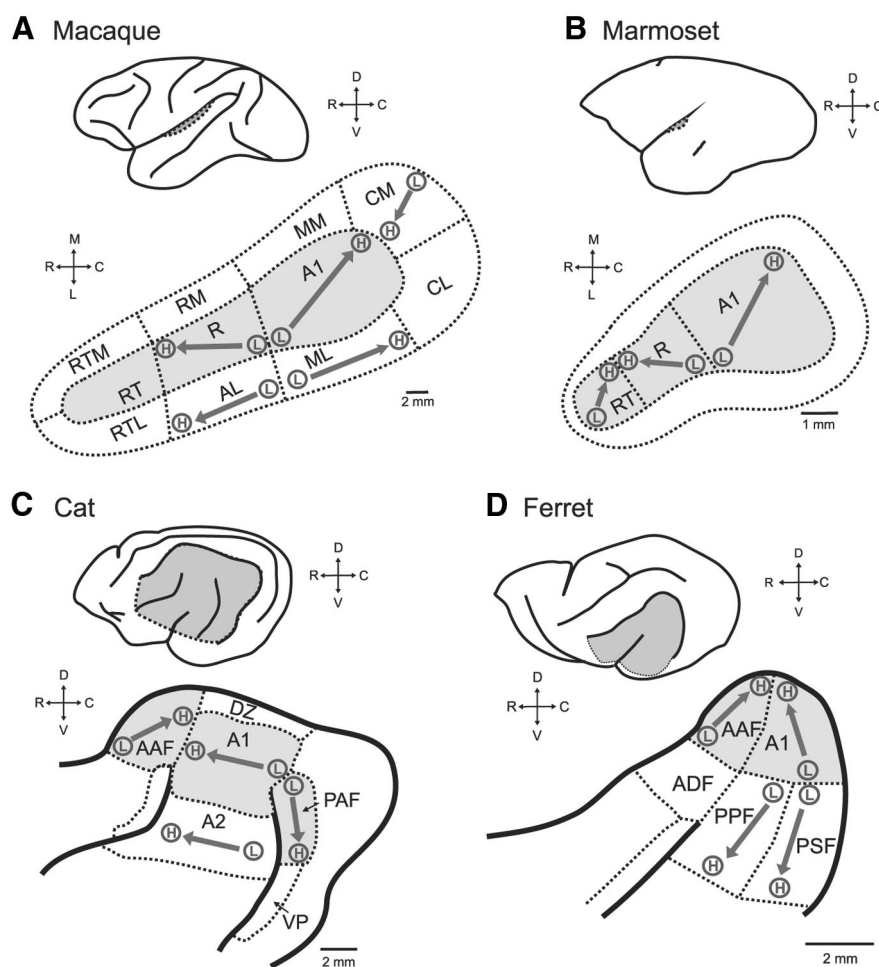


Figure 1. Schematics of auditory cortex across four species: **A**, Macaque; **B**, marmoset; **C**, cat; and **D**, ferret. In each panel, the top schematic shows an outline of the brain with auditory cortex indicated (dotted line, gray). The bottom schematic in each panel shows a closer view of the auditory cortex, with sulci (solid lines) and field boundaries (dotted lines). Core auditory cortex is shaded gray, and the orientations of known tonotopic maps are indicated with an arrow from low (L) to high (H) frequencies. A1, Primary auditory cortex; A2, secondary auditory area; AAF, anterior auditory field; ADF, anterior dorsal field; AL, anterolateral belt; CL, caudolateral belt; CM, caudomedial belt; DZ, dorsal zone; ML, mediolateral belt; MM, mediomedial belt; PAF, posterior auditory field; PPF, posterior pseudosylvian field; PSF, posterior suprasylvian field; r, rostral field; RM, rostromedial belt; RT, rostral temporal field; RTL, rostromedial lateral belt; RTM, rostromedial medial belt; VP, ventral posterior field.

neurons, but not single neurons, provided sufficient F0 discrimination to account for ferrets' thresholds on an equivalent behavioral task, in which ferrets were trained to categorize the same sounds as "low" or "high." This group further showed that although neurons that were sensitive to the periodicity of artificial vowels could be found across 5 examined fields of primary and secondary auditory cortex, these neurons were not selective (i.e., specialized) for pitch (Bizley et al., 2009). A neuron that was sensitive to vowel pitch almost always carried information about the timbre or spatial location of the vowel as well. For a subset of these neurons, feature representations were multiplexed within separate response time windows, so the pitch and timbre of vowels can be invariantly represented in a single auditory cortical neuron (Walker et al., 2011). It may therefore be instructive for future studies to examine pitch tuning in multiple time windows throughout a neuron's response to sounds, since tuning properties across the onset, sustained, and offset windows can be fundamentally different (Wang et al. 2005).

Given that humans and animals do experience a percept of pitch that generalizes across a variety of sounds with the same periodicity (including missing fundamental sounds), it seems

reasonable to expect to find neurons at some level of the auditory system that integrate the periodicity cues described above to compute a stimulus-invariant pitch representation. Such neurons may be distributed throughout the auditory cortex, but human imaging studies suggest that pitch neurons are likely to be concentrated in a region of auditory cortex that is specialized to encode the periodicity of temporally regular sounds (Griffiths and Hall, 2012). Bendor and Wang (2005) have provided evidence for pitch extraction by single neurons in the auditory cortex of awake marmosets. “Pitch-selective neurons” described in this study were defined as those whose CF for pure tones matched their periodicity tuning for missing fundamental harmonic complex sounds, where the spectral components of the latter all lay outside of the neuron’s excitatory-frequency response area. The region containing the pitch-selective neurons is confined to the low-frequency border of A1, R, and lateral belt areas. Approximately 39% of neurons within this anterolateral pitch region were classified as pitch-selective neurons based on multiple criteria. Using temporally jittered click trains and iterated rippled noises, Bendor and Wang (2010) further demonstrated that these pitch-selective neurons were sensitive to the temporal regularity of sounds, unlike modulation-sensitive neurons outside the pitch area, which are instead tuned to repetition rate regardless of the waveform’s temporal regularity. A recent study has reported evidence of pitch perception by marmosets (Osmanski et al., 2011).

A cortical region analogous to the pitch region reported by Bendor and Wang (2005) has been identified in several human imaging studies, but has yet to be identified in other animal species for which behavioral evidence of pitch perception is established. But if it does exist, one shall expect to find it in similar low-frequency borders of primary and non-primary cortical fields. While such a region was not specifically investigated in the study by Bizley et al. (2009), that study observed an overrepresentation of neurons sensitive to the pitch of artificial vowels in the low-frequency borders of A1 and tonotopic secondary fields in the ferret. Ultimately, lesion or cortical inactivation studies will be a necessary complement to these single-neuron investigations to establish the behavioral relevance of putative pitch codes.

Outstanding questions and directions for future work

There are several key questions that remain to be addressed in future studies of pitch encoding in auditory cortex. First, is there more than one “pitch center” in auditory cortex? If so, what are specific roles played by these different pitch-processing centers? To answer these questions, researchers need to pursue investigations to identify cortical regions, in particular in the secondary auditory cortex, that extract pitch embedded in complex sounds using a wide range of stimuli. Second, what are underlying mechanisms that allow any “pitch neuron” to perform pitch extraction computations? To understand such questions, one has to employ techniques beyond extracellular recordings, such as intracellular and two-photon optical recordings. Third, is a putative pitch region involved in pitch perception? While a demonstration of parallel properties between an animal’s behavioral responses and corresponding neural responses is a useful step toward answering such a question, a more convincing demonstration would be to show that the interruption of neural activity in a putative pitch region will lead to alterations in an animal’s pitch perception performance. Newly emerged optogenetic techniques could be a useful tool in this line of research in addition to other inactivation methods.

Computationally, it is important to differentiate between neurons (or cortical regions) that extract pitch embedded in complex sounds (such as harmonic complex) and those that bear pitch information. This requires examining whether pitch is specifically or uniquely represented by the neuron or cortical region under study, as well as determining that the physiological signal of the neurons corresponds with the animal’s perception of pitch. Neural responses bearing pitch information or encoding acoustic parameters associated with pitch can be found throughout much of the ascending auditory pathway (Cariani and Delgutte, 1996a,b), though their specificity for pitch may increase at successive higher processing stages. Neural representations earlier in the system could serve as precursors to the neurons that ultimately compute pitch, but they may not represent the final stages of pitch processing. Technically, it is crucial to apply rigorous controls to rule out influences by such factors as acoustic artifacts and cochlear distortions before a neuron or cortical region is considered “pitch-selective.”

References

- Bendor D, Wang X (2005) The neuronal representation of pitch in primate auditory cortex. *Nature* 436:1161–1165. [CrossRef Medline](#)
- Bendor D, Wang X (2006) Cortical representations of pitch in monkeys and humans. *Curr Opin Neurobiol* 16:391–399. [CrossRef Medline](#)
- Bendor D, Wang X (2010) Neural coding of periodicity in marmoset auditory cortex. *J Neurophysiol* 103:1809–1822. [CrossRef Medline](#)
- Bizley JK, Nodal FR, Nelken I, King AJ (2005) Functional organization of ferret auditory cortex. *Cereb Cortex* 15:1637–1653. [CrossRef Medline](#)
- Bizley JK, Walker KM, Silverman BW, King AJ, Schnupp JW (2009) Interdependent encoding of pitch, timbre and spatial location in auditory cortex. *J Neurosci* 29:2064–2075. [CrossRef Medline](#)
- Bizley JK, Walker KM, King AJ, Schnupp JW (2010) Neural ensemble codes for stimulus periodicity in auditory cortex. *J Neurosci* 30:5078–5091. [CrossRef Medline](#)
- Cariani PA, Delgutte B (1996a) Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J Neurophysiol* 76:1698–1716. [Medline](#)
- Cariani PA, Delgutte B (1996b) Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *J Neurophysiol* 76:1717–1734. [Medline](#)
- Fishman YI, Reser DH, Arezzo JC, Steinschneider M (1998) Pitch vs. spectral encoding of harmonic complex tones in primary auditory cortex of the awake monkey. *Brain Res* 786:18–30. [CrossRef Medline](#)
- Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2000) Complex tone processing in primary auditory cortex of the awake monkey. II. Pitch versus critical band representation. *J Acoust Soc Am* 108:247–262. [CrossRef Medline](#)
- Griffiths TD, Hall DA (2012) Mapping pitch representation in neural ensembles with fMRI. *J Neurosci* 32:13343–13347. [CrossRef](#)
- Hackett TA, Preuss TM, Kaas JH (2001) Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol* 441:197–222. [CrossRef Medline](#)
- Heffner HH, Whitfield IC (1976) Perception of the missing fundamental by cats. *J Acoust Soc Am* 51:520–529.
- Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A* 97:11793–11799. [CrossRef Medline](#)
- Kadia SC, Wang X (2003) Spectral integration in A1 of awake primates: neurons with single and multi-peaked tuning characteristics. *J Neurophysiol* 89:1603–1622. [Medline](#)
- Kalluri S, Depireux DA, Shamma SA (2008) Perception and cortical neural coding of harmonic fusion in ferrets. *J Acoust Soc Am* 123:2701–2716. [CrossRef Medline](#)
- Langner G, Dinse HR, Godde B (2009) A map of periodicity orthogonal to frequency representation in the cat auditory cortex. *Front Integr Neurosci* 3:27. [Medline](#)
- Nelken I, Bizley JK, Nodal FR, Ahmed B, King AJ, Schnupp JW (2008) Re-

- sponses of auditory cortex to complex stimuli: functional organization revealed using intrinsic optical signals. *J Neurophysiol* 99:1928–1941. [CrossRef Medline](#)
- Osmanski MS, Bendor DA, Wang X (2011) Dual mechanisms for processing pitch in marmosets (*Callithrix jacchus*). *Soc Neurosci Abstr* 38:691.08.
- Oxenham AJ (2012) Pitch perception. *J Neurosci* 32:13335–13338. [CrossRef](#)
- Schulze H, Langner G (1997) Periodicity coding in the primary auditory cortex of the Mongolian gerbil (*Meriones unguiculatus*): two different coding strategies for pitch and rhythm? *J Comp Physiol A* 181:651–663. [CrossRef Medline](#)
- Schulze H, Hess A, Ohl FW, Scheich H (2002) Superposition of horseshoe-like periodicity and linear tonotopic maps in auditory cortex of the Mongolian gerbil. *Eur J Neurosci* 15:1077–1084. [CrossRef Medline](#)
- Schwarz DW, Tomlinson RW (1990) Spectral response patterns of auditory cortex neurons to harmonic complex tones in alert monkey (*Macaca mulatta*). *J Neurophysiol* 64:282–298. [Medline](#)
- Steinschneider M, Reser DH, Fishman YI, Schroeder CE, Arezzo JC (1998) Click train encoding in primary auditory cortex of the awake monkey: evidence for two mechanisms subserving pitch perception. *J Acoust Soc Am* 104:2935–2955. [CrossRef Medline](#)
- Sutter ML, Schreiner CE (1991) Physiology and topography of neurons with multi-peaked tuning curves in cat primary auditory cortex. *J Neurophysiol* 65:1207–1226. [Medline](#)
- Tomlinson RW, Schwarz DW (1988) Perception of the missing fundamental in nonhuman primates. *J Acoust Soc Am* 84:560–565. [CrossRef Medline](#)
- Walker KM, Bizley JK, King AJ, Schnupp JW (2011) Multiplexed and robust representations of sound features in auditory cortex. *J Neurosci* 31:14565–14576. [CrossRef Medline](#)
- Wallace MN, Roeda D, Harper MS (1997) Deoxyglucose uptake in the ferret auditory cortex. *Exp Brain Res* 117:488–500. [CrossRef Medline](#)
- Wiegand L, Patterson RD (1999) Quantifying the distortion products generated by SAM noise. *J Acoust Soc Am* 106:2709–2718. [CrossRef Medline](#)
- Winer JA, Lee CC (2007) The distributed auditory cortex. *Hear Res* 229:3–13. [CrossRef Medline](#)