

Jan W. H. Schnupp and Jennifer K. Bizley

J Neurophysiol 103:1708-1711, 2010. doi:10.1152/jn.00182.2010

You might find this additional information useful...

This article cites 23 articles, 15 of which you can access free at:

<http://jn.physiology.org/cgi/content/full/103/4/1708#BIBL>

Medline items on this article's topics can be found at <http://highwire.stanford.edu/lists/artbytopic.dtl> on the following topics:

Physics .. Sound
Veterinary Science .. Auditory Cortex
Physiology .. Whales
Physiology .. Humans

Updated information and services including high-resolution figures, can be found at:

<http://jn.physiology.org/cgi/content/full/103/4/1708>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of April 25, 2010 .

On Pitch, the Ear and the Brain of the Beholder. Focus on “Neural Coding of Periodicity in Marmoset Auditory Cortex.”

Jan W. H. Schnupp and Jennifer K. Bizley

Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, United Kingdom

The pitch of a sound is perhaps one of its most salient properties. It is the pitch of sounds that allows us to distinguish and recognize melodies, and the salience of pitch and melody appears to be exploited by a strikingly wide variety of vertebrate life-forms, from song birds to “singing” hump-back whales all the way to human composers of advertising jingles. But while perceptually pitch has a great “immediacy,” physiologically pitch is a surprisingly complex phenomenon. It is sometimes said that the pitch of a sound is “related to its frequency content” but that relationship is anything but straightforward (Plack and Oxenham 2005). Many natural sounds contain large numbers of frequency components, and it is possible to produce an infinite variety of sounds with rather different frequency composition that nevertheless share the same pitch. Not all sounds evoke a clear pitch, but one thing that all sounds that do evoke a pitch percept have in common is that they are *periodic*, i.e., the waveform of the sound consists of a short “motif” that is repeated very rapidly over and over again, and the “speed” at which these repeats occur is the chief determinant of the perceived pitch value. Note that these repeated motifs are too brief to be perceived as distinct events. For humans to hear a clear musical pitch, their period must be roughly between 25 and 0.33 ms long (corresponding to pitches between ca. 40 and 3,000 Hz), and while the motifs needn’t be absolutely identical from one repeat to the next, they nevertheless have to be “similar.” If the pattern of repeated motifs becomes increasingly less regular, then the sound becomes increasingly noise-like, and the pitch becomes increasingly less salient until it becomes indistinguishable. Perceived pitch can therefore be thought of as a measure of the underlying “regularity” of a sound.

The requirement that sounds with a clear pitch must be periodic constrains their frequency content (i.e., their Fourier spectrum cannot take just any shape). A sound wave can only be periodic if all the sine wave components that make up the periodic sound “conform to the common underlying rhythm.” In other words, only sine waves that can fit a whole number of cycles into one period of the overall sound can be part of the spectrum of a periodic sound. Consequently, the spectra of periodic sounds are always composed of the harmonics (integer multiples) of a given “fundamental” frequency. So-called “place theories of pitch” assume that these harmonically spaced maxima in the spectrum of periodic sounds produce distinct peaks of excitation along the basilar membrane in the inner ear, and that the position and spacing of these peaks is then interpreted by the brain and determines the perceived pitch. However, the fact that periodic sounds must be composed of harmonically related frequencies does not constrain

them very tightly. For example, periodic sounds can vary greatly in how strongly the various harmonics are represented. In fact, while some periodic sounds that produce strong pitch percepts contain dozens or hundreds of harmonics within the audible frequency range, others contain only a very small number of harmonics, and pure tones, perhaps the “archetypal” periodic sounds, contain only a single one.

Usually periodic sounds which have a lot of energy at higher harmonics sound “brighter” than those that do not, but while changing the relative amplitude of high and low harmonics can dramatically change the sound’s “timbre,” it usually does not change its pitch. In fact, in sounds with several prominent higher harmonics one can even reduce the amplitude of the lowest harmonic, the fundamental, to nothing, and while such “missing fundamental” stimuli will sound “less rich” than equivalent sounds that do carry significant sound energy at the fundamental, their pitch still remains the same. And one need not stop at removing only the fundamental; harmonic complexes missing not several of the lowest harmonics may still evoke a recognizable pitch at the missing fundamental.

Place theories of pitch have a hard time explaining how the perceived pitch can remain stable in the face of such radical alterations of a sound’s spectrum. Alternative explanations for the basis of pitch known as “timing theories” are therefore gaining widespread acceptance. These propose that our brains derive pitch mostly from “time domain” cues that are extracted from the phase locked temporal firing patterns of auditory nerve fibers (Cariani and Delgutte 1996a,b). But whether pitch information enters the auditory system mostly through temporally or through place-coded information, much additional neural processing is clearly required in either case before a clear and stable pitch percept can emerge.

The neural structures and mechanisms that underpin our pitch perception are still only poorly understood and are the subject of much active research. Indeed in two recent papers, Bendor and Wang have described discoveries made during recordings from the auditory cortex of the marmoset that shed new light on this issue. Their first paper (Bendor and Wang 2005) documented neurons that appeared tuned to the periodicity of harmonic tone complexes, even if these tone complexes were “missing fundamental” stimuli. The responses of these neurons are therefore apparently stable when the sound spectrum changes, as long as the periodicity of the sound remains constant, just as the perceived pitch remains constant under identical manipulations. Neurons exhibiting this stable periodicity tuning seemed to cluster anatomically in a location near the low frequency border of primary auditory cortex, an area that Bendor and Wang have therefore designated as a putative “pitch area.” In a second recent paper, Bendor and Wang (2010) go on to show that these putative “pitch neurons” not only respond to missing fundamental stimuli but appear tuned

Address for reprint requests and other correspondence: J. Schnupp, Dept. of Physiology, Anatomy and Genetics, Sherrington Building, University of Oxford, Parks Road, Oxford OX1 3PT, UK (E-mail: jan.schnupp@dpag.ox.ac.uk).

to the same specific periodicities across a range of different types of spectrally dissimilar periodic test stimuli. In other words, sounds that elicit a consistent pitch percept in human listeners evoke equivalent responses in a subset of neurons within the putative pitch area. Furthermore, and perhaps most interestingly, the strength of these responses declines if originally highly periodic stimuli (regular pulse trains) are rendered increasingly less periodic by the introduction of “temporal jitter” in the stimulus period. This decline in response strength appears to mimic the decline in pitch salience observed when we listen to such stimuli.

What makes the results presented by Bendor and Wang in these papers so exciting is that the responses they observed in these putative pitch neurons seem to mirror some of the defining psychophysical measures of pitch perception in humans. These parallels strongly suggest that these neurons may have a key role to play in pitch perception, but it is important not to jump to conclusions. Bendor and Wang classify a neuron as a pitch neuron if it fulfils a list of criteria, which are essentially designed to ascertain three facts: first, that the neuron is “tuned” to a defined range of stimulus periodicities, second, that this tuning persists regardless of whether the stimulus is a pure tone, a harmonic complex or an iterated rippled noise (i.e., a noise that has been rendered periodic by an iterated delay and add operation), and third, that the neuron’s responses decline if the stimulus becomes less periodic. The discovery of neurons the responses of which meet these rather strict criteria is in itself remarkable because most neurons within the central auditory pathway appear to be far more interested in the sound’s frequency content than its periodicity. However, it is worth stepping back for a moment to ask whether the criteria adopted by Bendor and Wang (2005, 2010) are really necessary and sufficient to define a pitch neuron. In this context, we would like to raise three points of discussion. First we want to consider the assumptions that are implicit in the proposed physiological definition a pitch neuron. Second we want to ask more generally what criteria a neural substrate for pitch perception must necessarily fulfill. And finally, we ask to whether any physiological stimulus-response criteria can ever serve as a sufficient definition of a psychological dimension.

One of the key questions in the sensory neurosciences concerns which features of the neural code are “read-out” to support perception. Bendor and Wang’s criteria for pitch neurons imply a requirement for “tuning,” which here is taken to mean that the firing rate as a function of stimulus pitch must be nonmonotonic, with a single peak, and that only a limited pitch range should produce a response. While sensory neurons in many modalities indeed exhibit single peaked tuning curves for all manner of stimulus features, from sound frequency tuning in the auditory nerve to orientation tuning in primary visual cortex, there is nevertheless no a priori reason to assume that all stimulus properties must necessarily be represented in that way. Consider neurons in the lateral superior olive (LSO) that encode interaural level differences (ILDs), one of the main cues to the horizontal location of a sound source. The LSO does not employ sets of neurons that are “tuned” to limited ranges of ILD. Instead, LSO neurons exhibit a sort of “population rate code,” as LSO neurons typically show monotonic increases of their firing rate as a function of ILD through the entire range of physiological ILDs

(Irvine et al. 2001). For neurons in auditory cortex, the relationship between neural firing pattern and spatial location of a sound source is often complex and seemingly arbitrary, but it is still possible to extract large amounts of information about the spatial location from the neural response (Middlebrooks et al. 1998; Morsic-Flogel et al. 2005). From an information coding point of view, the “shape of the tuning curve” is somewhat secondary. Much more important is whether there is a reliable statistical interdependence between particular neural firing patterns and particular stimulus parameters (Nelken et al. 2005). The fact that the pitch neurons of Bendor and Wang (2005, 2010) appear tuned to a preferred “best fundamental frequency” is very interesting, but it is by no means certain that all neurons that support pitch perception must necessarily be tuned in such a manner.

A second key aspect of Bendor and Wang’s (2005, 2010) criteria for pitch neurons is based around the assumption that the pitch constancy that we observe when the spectral profile of a periodic sound changes must be reflected directly in an equivalent constancy of the response of individual pitch neurons. The logic of this argument is, in many respects, hard to fault, given that pitch constancy is indeed one of the defining characteristics of pitch. Throughout musical traditions one can find many examples where a particular melodic theme is set up by one particular instrument or voice, only to be taken up later by another. Such artistic devices are the bread and butter of any skilled musical composer, but they would be entirely lost on us if pitch was not to a large degree independent of the spectral timbre of a particular sound source. However, pitch sensitivity and pitch perception most likely evolved originally for more prosaic purposes than the appreciation of musical composition. One such unmusical, but nonetheless highly useful feature of pitch is that it can serve as a powerful cue in auditory scene analysis. We find it much easier to separate a foreground sound from a background if the two sound sources that need to be segregated differ in pitch (De Cheveigne et al. 1997a,b). However, if we were trying to build a device that used pitch as one of a number of acoustic cues to facilitate acoustic scene segregation, then we would most likely find it preferable to base such a device on components that were sensitive to particular combinations of pitch and spectral timbre rather than on components that are sensitive to pitch alone, irrespective of other acoustic features. Many sound sources will, after all exhibit characteristic joint distributions of pitch and timbre. Perhaps that explains why the responses of the overwhelming majority of auditory cortical neurons do not exhibit pitch constancy but appear instead to be jointly sensitive to stimulus periodicity as well as spectral timbre and sound source location (Bizley et al. 2009).

The apparent pitch constancy of the physiological responses described by Bendor and Wang (2005, 2010) is intriguing in large part because it is so rare—indeed several neuroimaging studies failed to find spectrally invariant “maps” of stimulus periodicity in cortex (Hall and Plack 2007; Nelken et al. 2008). Many neurons that fail to meet Bendor and Wang’s “physiological pitch constancy criterion” may nonetheless incorporate pitch into their operations, but one could argue that presumably those that do meet the criterion are in some sense “closer” to the psychological perceptual phenomenon of pitch. On the other hand, it is in principle possible that pitch constancy could be an emergent property of a cortical network which is com-

posed of neurons that individually do not exhibit pitch invariant responses. Whether the small number of pitch neurons that Bendor and Wang (2005, 2010) have identified really have a privileged or causal role to play in the generation of pitch percepts can only be decided with further experiments. For example, it would be of great interest to know whether inactivating or disrupting the activity of the putative pitch area produced a pitch perception deficit.

That of course raises a wider question regarding how future studies of the physiological basis of pitch perception ought to be designed. Bendor and Wang argue that, if it is possible to identify “face cells” in higher order visual cortex based on particular stimulus-response relationships (Tsao et al. 2006), then it ought to be possible similarly possible to identify pitch cells in auditory cortex using a suitably chosen set of stimulus-response criteria. However, faces and pitch differ in a quite fundamental way: a face remains a face when nobody looks at it. Faces are physical objects, and it may well be possible to capture their essential characteristics in a suitably chosen external stimulus set. The pitch of a sound, however, has no independent physical existence. Like pain or color, pitch is quintessentially a psychological phenomenon. A sound that nobody hears has no pitch at all. While faces are recognized by the brain, pitches are created in the brain, and charting the relationship between physical stimulus properties and responses in auditory cortex is therefore only telling half the story. The crucial link between neural activity and the psychological percept remain unexplored.

Perceived pitch and stimulus periodicity are of course usually very tightly correlated, and one might argue that their coupling is so tight that a neuron that is selective for stimulus periodicity *must* be selective for pitch. But let us consider this briefly in the light of the phenomenon of tinnitus. Tinnitus is a pathological condition in which patients report a clear auditory percept (“a ringing in the ear”) that occurs, often persistently, even in the absence of any kind of acoustic stimulation. Furthermore, in the large majority of cases, tinnitus has a clear and identifiable pitch. Tinnitus is not a rare phenomenon. Most people have experienced it very occasionally, and ~1 in every 200 people suffer from it chronically. Thus in the U.S. alone, over 10 million people have almost constant pitch perceptions that are completely unrelated to any acoustic stimulus properties. Clearly, a full understanding of the neural basis of pitch as a psychological, or indeed as a physiological, phenomenon cannot be based on stimulus-response relationships alone.

Much remains to be done to confirm and elucidate which role the pitch neurons discovered by Bendor and Wang (2005, 2010) really do have to play in pitch perception. Attempts to clarify the relationship between neural activity and subjective perception have a long tradition in the visual neurosciences. Some of the earliest methods deployed are the so-called “neurometric” approaches that look for global parallels between putative neural codes and perceptual discrimination ability as measured by psychometric tests (Tolhurst et al. 1983). These were later followed by “choice probability” measures that try to determine how much a neural response might contribute to an animal’s percept of a stimulus on a trial-by-trial basis (Celebrini and Newsome 1994; Krug et al. 2004), and there were even some successful attempts to bias an animal’s perception by manipulating neural responses through electrical stimulation (Salzman and Newsome 1994). Each of these

methods has its advantages as well as its limitations (Krug 2004), but between them they can provide considerable insight into the likely neural basis of perception. The auditory neurosciences have, however, been slow to adopt such methodology. Neurometric techniques have only recently been introduced (Lemus et al. 2009; Walker et al. 2008; Wang et al. 2007), and more advanced techniques, such as choice probability calculations (Russ et al. 2008) are also only slowly being adopted. Such approaches remain very much the exception in auditory neuroscience, probably because too few auditory laboratories are set up to carry out behavioral testing of perceptual performance and electrophysiological recording in parallel. Technically, such studies are extremely demanding. The auditory neuroscience community therefore has a lot of catching-up to do before we will be able to say with confidence whether the periodicity tuned neurons described by Bendor and Wang (2010) truly deserve their designation as pitch neurons. Of course, none of these wider interpretational issues detract from the fact that the experiments described by Bendor and Wang (2010) in this issue are beautiful, have yielded very valuable and informative data, and represent an important step forward in our journey toward a deeper understanding of the neural basis of auditory perception.

REFERENCES

- Bendor D, Wang X. The neuronal representation of pitch in primate auditory cortex. *Nature* 436: 1161–1165, 2005.
- Bendor D, Wang X. Neural coding of periodicity in marmoset auditory cortex. *J Neurophysiol* 103: 1809–1822, 2010.
- Bizley JK, Walker KM, Silverman BW, King AJ, Schnupp JW. Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. *J Neurosci* 29: 2064–2075, 2009.
- Cariani PA, Delgutte B. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J Neurophysiol* 76: 1698–1716, 1996a.
- Cariani PA, Delgutte B. 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, 1996b.
- Celebrini S, Newsome WT. Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J Neurosci* 14: 4109–4124, 1994.
- De Cheveigne A, Kawahara H, Tsuzaki M, Aikawa K. Concurrent vowel identification. I. Effects of relative amplitude and F_0 difference. *J Acoust Soc Am* 101: 2839–2847, 1997a.
- De Cheveigne A, McAdams S, Marin CMH. Concurrent vowel identification. II. Effects of phase, harmonicity, and task. *J Acoust Soc Am* 101: 2848–2856, 1997b.
- Hall DA, Plack CJ. The human “pitch center” responds differently to iterated noise and Huggins pitch. *Neuroreport* 18: 323–327, 2007.
- Irvine DR, Park VN, McCormick L. Mechanisms underlying the sensitivity of neurons in the lateral superior olive to interaural intensity differences. *J Neurophysiol* 86: 2647–2666, 2001.
- Krug K. A common neuronal code for perceptual processes in visual cortex? Comparing choice and attentional correlates in V5/MT. *Phil Trans R Soc Lond* 359: 929–941, 2004.
- Krug K, Cumming BG, Parker AJ. Comparing perceptual signals of single V5/MT neurons in two binocular depth tasks. *J Neurophysiol* 92: 1586–1596, 2004.
- Lemus L, Hernandez A, Romo R. Neural codes for perceptual discrimination of acoustic flutter in the primate auditory cortex. *Proc Natl Acad Sci USA* 106: 9471–9476, 2009.
- Middlebrooks JC, Xu L, Eddins AC, Green DM. Codes for sound-source location in nontopographic auditory cortex. *J Neurophysiol* 80: 863–881, 1998.
- Mrsic-Flogel TD, King AJ, Schnupp JW. Encoding of virtual acoustic space stimuli by neurons in ferret primary auditory cortex. *J Neurophysiol* 93: 3489–3503, 2005.

- Nelken I, Bizley JK, Nodal FR, Ahmed B, King AJ, Schnupp JW.** Responses of auditory cortex to complex stimuli: functional organization revealed using intrinsic optical signals. *J Neurophysiol* 99: 1928–1941, 2008.
- Nelken I, Chechik G, Mrsic-Flogel TD, King AJ, Schnupp JW.** Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. *J Comput Neurosci* 19: 199–221, 2005.
- Plack CJ, Oxenham AJ.** The psychophysics of pitch. In: *Pitch: Neural Coding and Perception*, edited by Plack CJ, Oxenham AJ, Fay RR, Popper AN. New York: Springer, 2005, p. 7–55.
- Russ BE, Orr LE, Cohen YE.** Prefrontal neurons predict choices during an auditory same-different task. *Curr Biol* 18: 1483–1488, 2008.
- Salzman CD, Newsome WT.** Neural mechanisms for forming a perceptual decision. *Science* 264: 231–237, 1994.
- Tolhurst DJ, Movshon JA, Dean AF.** The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res* 23: 775–785, 1983.
- Tsao DY, Freiwald WA, Tootell RB, Livingstone MS.** A cortical region consisting entirely of face-selective cells. *Science* 311: 670–674, 2006.
- Walker KM, Ahmed B, Schnupp JW.** Linking cortical spike pattern codes to auditory perception. *J Cogn Neurosci* 20: 135–152, 2008.
- Wang L, Narayan R, Grana G, Shamir M, Sen K.** Cortical discrimination of complex natural stimuli: can single neurons match behavior? *J Neurosci* 27: 582–589, 2007.