Neurons and objects: the case of auditory cortex

Israel Nelken[1,2,†] *and Omer Bar-Yosef*[3]

[1] Department of Neurobiology, The Silberman Institute of Life Sciences, Edmund Safra Campus, Hebrew University, Jerusalem, Israel.
[2] The Interdisciplinary Center for Neural Computation, Hebrew University, Jerusalem, Israel.
[3] Department of Pediatrics, Safra Children's Hospital, Sheba Medical Center, Tel Hashomer, Israel.

Sounds are encoded into electrical activity in the inner ear, where they are represented (roughly) as patterns of energy in narrow frequency bands. However, sounds are perceived in terms of their high-order properties. It is generally believed that this transformation is performed along the auditory hierarchy, with low-level physical cues computed at early stages of the auditory system and high-level abstract qualities at high-order cortical areas. The functional position of primary auditory cortex (A1) in this scheme is unclear – is it 'early', encoding physical cues, or is it 'late', already encoding abstract qualities? Here we argue that neurons in cat A1 show sensitivity to high-level features of sounds. In particular, these neurons may already show sensitivity to 'auditory objects'. The evidence for this claim comes from studies in which individual sounds are presented singly and in mixtures. Many neurons in cat A1 respond to mixtures in the same way they respond to one of the individual components of the mixture, and in many cases neurons may respond to a low-level component of the mixture rather than to the acoustically dominant one, even though the same neurons respond to the acoustically-dominant component when presented alone.

Keywords: auditory cortex, complex sounds, cats, electrophysiology, single neurons, auditory objects

THE AUDITORY SYSTEM AND COMPLEX SOUNDS

The auditory system uses a time-varying signal – air pressure at the ear drum – for unimaginably complex tasks, such as understanding speech, enjoying music or detecting very small changes in the soundscape that may indicate the approach of a hungry tiger. In order to perform all of these tasks, the mammalian auditory system has to accommodate two conflicting requirements. The first is the need to encode sounds in all their details. The highly detailed representation of sounds is required in order to be able to detect informative, but physically small and context-dependent, changes in the soundscape. Such a highly detailed representation of the incoming sound stream is however far too complex to guide behavior, since much of this detail is irrelevant in any given situation. The second, conflicting requirement is, therefore, the need to extract the relevant information out of the highly detailed representation of the sound stream, in order to encode sounds in terms that are useful for guiding future behavior.

Representations of the relevant information may collapse many physically-different sounds into the same class: for example, when we have to understand a spoken message, a high-fidelity recording of the message, and the same message heard through a cellular phone, would be considered 'the same'. There may be, however, many different possible categorizations of these sounds: if we have to judge the quality of the sound reproduction, the high-fidelity recording and the cellular phone message will be perceived as very good and very bad, respectively. Thus, in both tasks, 'understanding' and 'quality judgment', a large

set of physically-different sounds are classified into a substantially smaller number of classes that are behaviorally meaningful. In that sense, the behaviorally-relevant representations are coarse (because each class may contain many different physical sounds). This doesn't mean that extracting the class label in each case is easy – extracting the sequence of spoken words from the acoustic waveform is the notoriously difficult problem of automatic speech recognition, and quality judgments are very hard to capture by computer programs operating on the sounds themselves. Furthermore, sounds that have the same meaning may have very different quality, whereas sounds with the same quality may have very different meaning; thus, the labels in the 'understanding' and in the 'quality judgment' tasks are based on possibly very different physical features of the sounds, so that neither of them can fully replace the physically-detailed representation of the same sounds.

Thus, the auditory system has a 'double personality': it encodes sounds simultaneously with very high fidelity (Dallos, 1996) and also in a behaviorally-relevant, coarse way (Yost and Sheft, 1993). This double personality can be observed in perceptual experiments. On the one hand, humans can discriminate low-level physical features of sounds astoundingly well: the smallest frequency difference that can be perceived (admittedly, by highly-trained humans) is about 0.2% (whereas the basic interval of western music, the half-tone, is 6%, Moore, 2003). On the other hand, categorical perception experiments show that when a physical change in a speech sound is not informative, it will be not be perceived very well (Eggermont and Ponton, 2002).

The tension between high-resolution representations and coarse but behaviorally-relevant representations is reflected in current thinking about the structure and function of the auditory system. The representation of sounds at the level of the auditory nerve, the cochlear nucleus (the first central station of the auditory system), and the inferior colliculus (an obligatory midbrain station without an obvious homolog in other sensory systems), is generally considered to be highly detailed (e.g., Casseday et al., 2002). Single trials of single neurons in inferior colliculus are sufficient to reach human discrimination levels when decoded optimally, admittedly with a lot of precautionary notes (e.g., Shackleton et al., 2003, 2005). In this sense, the inferior colliculus may be considered as the neural substrate underlying the HiFi industry. On the other hand, responses in primary auditory cortex (A1) of cats and rodents are much coarser, leading already many years ago to the suggestion that A1 encodes sounds in terms that go beyond their acoustic structure (e.g., Wollberg and Newman, 1972). Human fMRI

data seems to be roughly consistent with such a hierarchical model. For example, speech is processed along multiple streams originating in A1, with more abstract representations occurring farther away from this structure (Scott and Johnsrude, 2003). Similarly, the most detailed conceptualization of sound processing in the primate brain, the 'where and what pathways' model (Rauschecker and Tian, 2000; Romanski et al., 1999) and its more elaborated descendents (Belin and Zatorre, 2000; Zatorre et al., 2002), suggests that the acoustic information, which is encoded in physical terms in A1, might be processed along a posterior pathway to extract its spatial properties ('where') and along an anterior pathway to extract other high-level properties. In both models, A1 is 'low-level' while the coarser, behaviorally-relevant representations evolve along the information streams away from A1.

Evidence for or against abstract representations of sounds by single neurons in A1 has been elusive. By far, most studies of A1 have used, and still use, very simple stimuli such as pure tones and broad-band noise (Bizley et al., 2005; Hromadka et al., 2008; Moshitch et al., 2006; Nelken et al., 2004; Qin et al., 2003; Read et al., 2002; Tan et al., 2004; Ulanovsky et al., 2003; Wehr and Zador, 2003). There are very good reasons for these choices. The use of complex sounds requires complex experimental designs, not always compatible with the limited amount of time available for recording neural responses. Thus, for example, Wollberg's studies on the coding of species-specific vocalizations in the squirrel monkey A1 started with the speculation that neurons in A1 encode these sounds according to their meaning rather than according to their acoustic structure (Wollberg and Newman, 1972), but ended 20 years later with a paper suggesting population coding of acoustic features (Pelleg-Toiba and Wollberg, 1991). This change occurred to a large extent because more rigorous acoustic controls have been applied over the years (e.g., Glass and Wollberg, 1983). Complex sounds also require more complex data analysis tools, which were lacking in the past and are currently in active development (Ahrens et al., 2008; Nelken and Chechik, 2007). On the other hand, even pure tones evoke highly non-trivial response in auditory cortex. To mention just a few such results, the highly complex structure of binaural interactions (Semple and Kitzes, 1993), the high sensitivity of cortical neurons to sound onsets (Heil and Irvine, 1996; as interpreted by Fishbach et al., 2001; Heil and Neubauer, 2003), or sound-level independent frequency coding (Phillips and Orman, 1984) have all been crucial for our current understanding of auditory cortex function and were all described in studies using simple stimuli such as pure tones and broadband noise bursts. Nevertheless, these studies

could not address the issue of abstract representations in auditory cortex.

In fact, currently, even research programs that employ complex sounds to study A1 tend to avoid the issue of abstract representations. Instead, such programs emphasize the interpretation of the responses in terms of feature detection, where the features might be complex but relatively limited in time (with a typical duration of 10–100 ms). Along this line, the highly influential paper of deCharms et al. (1998) showed a selection of spectro-temporal receptive fields (STRF) with complex structure, suggesting that these are an important feature of sound processing in A1. However, further research, both in awake and in anaesthetized animals, has generally failed to find highly complex STRFs, at least in the majority of cases (Ahrens et al., 2008; Depireux et al., 2001). Using a different approach, Wang et al. (2005) demonstrated the presence, in A1 of awake marmosets, of neurons that can be tonically driven but only by a stimulus with rather specific set of physical parameters such as frequency, sound level, or amplitude and frequency modulation patterns, and rates.

While these data emphasize the view of A1 as a collection of highly specific feature detectors, all of the above studies still used artificial sounds (even if rather complex ones). It remains possible that an abstract representation in A1 will only be elicited by more natural sounds, varying along appropriate abstract dimensions. However, what are precisely those abstract dimensions? Most studies using species-specific vocalizations did not uncover any dramatically different responses to such sounds. More generally, natural sounds do not evoke responses that are very different from those evoked by artificial sounds in A1. The recent study of Hromadka et al. (2008), for example, showed that in A1 of awake rats, population responses to pure tones and to an ensemble of natural sounds are rather similar – if at all, the average responses to tones were somewhat larger than the responses to natural sounds.

CAT AUDITORY CORTEX AND BIRD CHIRPS

We would like to argue here that there is a common conceptual drawback to all of these attempts to study A1. In fact, all of the above studies address the 'easy problem' of auditory processing: that of discriminating sounds from each other (Nelken and Ahissar, 2006). At least some neurons in A1 should be able to discriminate between any tokens of different phonemes or different words: these are stimuli that are physically different, and therefore they evoke different patterns of activity in the auditory nerve (e.g., Hienz et al., 1996; May et al., 1996). Such differences should be detectable also by neural activity at the level of the auditory cortex. The 'hard problem' of

auditory processing is very different. It is the fact that in real life, sounds are rarely heard by themselves. Instead, we usually hear sounds in mixtures, and the task of separating a mixture of sounds into its individual components is a hard computational problem (Bregman, 1990; Wang and Brown, 2006). We claim that the solution to this problem is an abstract high-level representation, because the notion of a component of a sound mixture is a high-level one. Sounds can be decomposed into components in many different ways. A case in point is the decomposition of sounds into narrow frequency bands that is performed at the inner ear. This decomposition, while basic to the function of the auditory system, does not result in most cases in perceptually-relevant components. For example, a complex periodic sound will be broken into groups of harmonics that will be represented in different frequency channels, but none of these components will be perceptually present in the sound in any obvious way (except maybe to highly-trained listeners!).

Inspite of this indeterminacy, the solution to the problem of separating a mixture of sounds into perceptually-valid components is very often obvious in natural settings, and is related (although not identical) to the problem of separating a sound mixture into components that have been emitted by different sound sources. Such decompositions make ecological sense, since they represent the sound in terms that are useful for guiding behavior. These notions are obviously close to the highly charged term 'auditory object' (Griffiths and Warren, 2004; Winkler et al., 2006).

To be concrete, we are going to limit ourselves to a rather simple situation: that of a main, narrowband sound such as a bird chirp, heard as a part of a natural noisy scene, but dominating it. In this case, the auditory objects are obvious: the bird chirp is one object, while the background noise is another, with possible further subdivisions of the background. **Figure 1** illustrates a specific example, used by Bar-Yosef and Nelken (2007). Here the full natural sound (Natural chirp) is divided into the clean Main chirp and the background noise (Noise). The Noise has two distinct components: a narrowband component in the range of frequencies that are present in the chirp, presumably representing the echoes of the chirp; and a wideband component that may result from multiple sources. If the Main chirp is sufficiently louder than the background noise, it will be easy to detect. Thus the interesting auditory question in this case is whether there are neurons that encode specifically the background (Noise in **Figure 1**, or even the weaker background component), in spite of the fact that it is acoustically much weaker. But why should encoding the background be an important or interesting auditory task? One answer to this question is
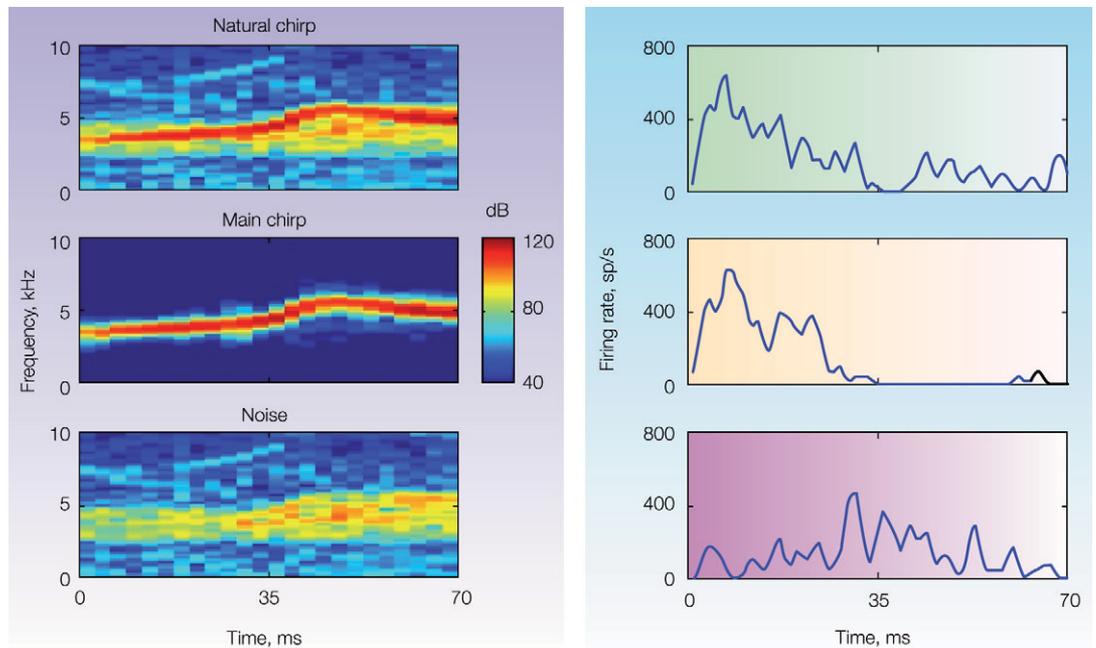
Figure 1 | **Spectrograms (left) and peri-stimulus time histograms of the responses of a primary-like unit in the ventral cochlear nucleus of a cat (right).** The stimuli were a segment of a natural sound consisting of a bird chirp and its natural background (top), the clean bird chirp (middle) and the background (bottom). These sounds have been used in Bar-Yosef and Nelken (2007).

that the background might contain highly pertinent information to a mid-size predator such as the cat: it may contain information regarding other prey (such as the mouse rustling in the grass) and about natural enemies (such as humans trying to catch it).

It is difficult to demonstrate that a neuron respond specifically to an acoustic component. To do so, it is necessary to contrast a standard representation of the neural responses, such as a frequency response area that charts the responses to pure tones, with the responses to the natural sounds. At the least, such a neuron should respond identically to all stimuli that contain this acoustic component, independent of other sounds that are simultaneously present. Furthermore, we would like to have neurons that respond as expected to pure tones, but that have surprising responses to the natural sounds and their components, responses that cannot be accounted for by the frequency response area but are consistent with the idea of the representation of abstract entities. The cochlear nucleus neuron whose responses are shown in **Figure 1** clearly does not show such specific responses: it respond to the natural chirp and to the main chirp with similar response patterns, while it responds to the noise with a different response pattern.

In two papers, Bar-Yosef and colleagues (Bar-Yosef and Nelken, 2007; Bar-Yosef et al., 2002) demonstrated the presence of neurons that show such specific responses to acoustic components in cat A1.

Figure 2 shows an example of the responses of such a neuron, and the way that the neural responses to a natural sound and its components were studied. The frequency response area of the neuron is shown at the top **Figure 2**. The plot shows response (coded by color) as a function of frequency and sound level. The power spectra of the clean bird chirp (white) and the background (red) are superimposed, to emphasize the fact that this neuron is sensitive to the frequencies present in the clean bird chirp, and that the chirp is substantially louder than the background. Each of the panels below is composed of the spectrogram of a stimulus, together with a raster plot showing the spike patterns evoked by 20 repetitions of that sound. The process of decomposition of the sound is illustrated by the tree structure: at the top, the response to the full natural sound is shown. The neuron responded to this sound with a vigorous onset response followed by a weaker response. The natural sound was separated into the clean chirp (middle row, right) and to the background (middle row, left). Surprisingly, the response to the clean chirp was rather weak, whereas the robust onset response was present in the response to the background. At a second step, the background itself was separated into echoes of the clean chirp (bottom row, right) and the wideband surround (bottom row, left). Again, surprisingly, the robust onset response was evoked by the low-level wideband surround.

Figure 2 | **The frequency response area of a neuron in cat primary auditory cortex (top) and its responses to a natural chirp and some of its modified versions.** Further details in the text.

alone) suggests that the onset response was evoked by the onset of the background, rather than by the onset of the main chirp, even when played together as part of the full natural sound. Thus, the neuron seems to have responded specifically to the wideband background, independently of presence of other, dominant acoustic component. Such capture of the responses of a neuron by a specific component of the natural sounds was typical (Bar-Yosef and Nelken, 2007). We would like to conclude that the capture of the responses of neurons in cat A1 by specific sound components is a partial solution of the problem of separating a sound mixture into its components. Furthermore, it seems that A1 neurons tend to perform the hard part of the decomposition process, in that they respond to the weak component in the mixture.

One possible critique of this interpretation of the results of Bar-Yosef and collaborators (2002, 2007) is that neurons in cat A1 might be especially sensitive to broadband sounds, and that this is the reason why neurons responded to the broadband components of the mixture. However, the neurons studied in these papers are rather typical to cat A1 (Bar-Yosef et al., 2002; Moshitch et al., 2006). Furthermore, Las et al. (2005) demonstrated a kind of converse to the results of Bar-Yosef and collaborators (2002, 2007). In their study, the loud acoustic component was fluctuating noise (such as is often found in a natural auditory scene, Nelken et al., 1999), while the low-level component was a pure tone. Las et al. (2005) showed that neurons in inferior colliculus (two synapses before auditory cortex) already showed extremely sensitive thresholds to tones in fluctuating noise (by responding to the tones in the energy minima of the fluctuating maskers), but that the responses of neurons in A1 were captured by the low-level tones, in a way that resembled the capture of similar A1 neurons by low-level wideband sound components as described by Bar-Yosef and Nelken (2007). Thus, neurons similar to those studied by Bar-Yosef and Nelken (2007) showed an extreme sensitivity to low-level tones (a narrowband stimulus) in a high-level masker (a wideband stimulus). It can be cautiously concluded that it is not a specific sensitivity of A1 neurons to wideband stimuli that is implicated here. As in the studies of Bar-Yosef and collaborators (2002, 2007), the results of Las et al. (2005) can be interpreted in terms of representation of auditory objects, except that here the low-level object is the pure tone, which is detected at the level of the inferior colliculus, but is represented much more explicitly and clearly in A1, in spite of its low level.

Support for the idea that neurons in A1 encode auditory objects is also given by another line of research: the highly sensitive stimulus-specific adaptation that was described in A1 of cats by

When tested with pure tones, the neuron in **Figure 2** was typical of A1, having a low-threshold narrow frequency response area (top of **Figure 2**). Neurons with such tuning curves in early auditory stations would respond to the bird chirp with or without the presence of the background noise (see **Figure 1** and Chechik et al., 2006). Cortical neurons with such narrow frequency response area would be expected to be insensitive to wideband stimuli (Schreiner and Mendelson, 1990). However, background noise had an inordinate influence on the responses of the neuron in **Figure 2**. In particular, the similarly robust onset response to the full natural sound and to the background (when presented

Ulanovsky et al. (2003, 2004). Oddball sequences, consisting of many repetitions of one pure tone frequency interspersed with rare presentations of another, nearby frequency, were used to study the effects of tone probability on single-neuron responses. Response size was inversely related to the probability tone presentation. The commonly presented tones showed adaptation with time constants on the order of 10–100 s, but rare tones that could be very close in frequency (10% or even 4% away from the common tone) showed much less or even no adaptation. This stimulus specificity was much better than the bandwidth of the neurons at the sound levels in which they were tested, which was an octave or more. Furthermore, behaviorally, frequency discrimination thresholds in cats are about 4% (Masterton et al., 1992). One interpretation of these results is that the neurons respond more strongly to the rare tone because it indicates the presence of a new auditory object in the soundscape.

CONCLUSIONS AND SPECULATIONS

What are the relationships between these findings and those showing high and complex selectivity of neural responses to physical cues, reviewed in section "The auditory system and complex sounds"? These two sets of findings are not necessarily in conflict. Many studies of parameter selectivity in A1 do not address in any way the issue of object representation and therefore cannot be used to refute the presence of object representation in A1. Furthermore, the high selectivity to specific complex physical features, (e.g., Wang et al., 2005) may create building blocks that give rise to object-related responses. Experiments like those described here, using natural or naturalistic mixtures, have not been conducted together with parametric studies of stimulus selectivity; such experiments are obviously necessary in order to elucidate the relationships between the two views of processing in auditory cortex discussed here.

How are we then to interpret the role of A1 in hearing? We suggest that A1 neurons encode much more than just the physical properties of recent sounds. In addition to these properties, they encode the context in which these sounds occur (as in Ulanovsky et al., 2003, 2004) and they may be highly selective to the auditory objects present in the scene (as in Bar-Yosef and Nelken, 2007; Bar-Yosef et al., 2002; Las et al., 2005). To the anatomical position of A1 as a hub from which information flows in different functional streams (Romanski et al., 1999), we can add now an important role in information processing. We would like to suggest that A1 produces representations of auditory objects, which are the entities that are processed in higher auditory areas. These higher auditory areas may assign properties such as pitch, spatial location, identity and meaning to the objects that are generated by the neural processes occurring in A1.

CONFLICT OF INTEREST STATEMENT

We 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.

ACKNOWLEDGEMENTS

REFERENCES

Ahrens, M. B., Linden, J. F., and Sahani, M. (2008). Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectrotemporal methods. *J. Neurosci.* 28, 1929–1942.

Bar-Yosef, O., and Nelken, I. (2007). The effects of background noise on the neural responses to natural sounds in cat primary auditory cortex. *Front. Comput. Neurosci.* 1.

Bar-Yosef, O., Rotman, Y., and Nelken, I. (2002). Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. *J. Neurosci.* 22, 8619–8632.

Belin, P., and Zatorre, R. J. (2000). 'What', 'where' and 'how' in auditory cortex. *Nat. Neurosci.* 3, 965–966.

Bizley, J. K., Nodal, F. R., Nelken, I., and King, A. J. (2005). Functional organization of ferret auditory cortex. *Cereb. Cortex* 15, 1637–1653.

Bregman, A. S. (1990). Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA, MIT Press.

Casseday, J. H., Fremouw, T., and Covey, E. (2002). The inferior colliculus: a hub for the central auditory system. In Integrative Functions in the Mammalian Auditory Pathway, Vol. 15, D. Oertel et al. eds (New York, NY, Springer), pp. 238–318.

Chechik, G., Anderson, M. J., Bar-Yosef, O., Young, E. D., Tishby, N., and Nelken, I. (2006). Reduction of information redundancy in the ascending auditory pathway. *Neuron* 51, 359–368.

Dallos, P. (1996). Overview: cochlear neurobiology. In The Cochlea, P. Dallos et al. eds (New York, NY, Springer), pp. 1–43.

deCharms, R. C., Blake, D. T., and Merzenich, M. M. (1998). Optimizing sound features for cortical neurons. *Science* 280, 1439–1443.

Depireux, D. A., Simon, J. Z., Klein, D. J., and Shamma, S. A. (2001). Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J. Neurophysiol.* 85, 1220–1234.

Eggermont, J. J., and Ponton, C. W. (2002). The neurophysiology of auditory perception: from single units to evoked potentials. *Audiol. Neurootol.* 7, 71–99.

Fishbach, A., Nelken, I., and Yeshurun, Y. (2001). Auditory edge detection: a neural model for physiological and psychoacoustical responses to amplitude transients. *J. Neurophysiol.* 85, 2303–2323.

Glass, I., and Wollberg, Z. (1983). Responses of cells in the auditory cortex of awake squirrel monkeys to normal and reversed species-specific vocalizations. *Hear. Res.* 9, 27–33.

Griffiths, T. D., and Warren, J. D. (2004). What is an auditory object? *Nat. Rev. Neurosci.* 5, 887–892.

Heil, P., and Irvine, D. R. F. (1996). On determinants of first-spike latency in auditory cortex. *Neuroreport* 7, 3073–3076.

Heil, P., and Neubauer, H. (2003). A unifying basis of auditory thresholds based on temporal summation. *Proc. Natl. Acad. Sci. USA* 100, 6151–6156.

Hienz, R. D., Aleszczyk, C. M., and May, B. J. (1996). Vowel discrimination in cats: thresholds for the detection of second formant changes in the vowel/epsilon. *J. Acoust. Soc. Am.* 100, 1052–1058.

Hromadka, T., Deweese, M. R., and Zador, A. M. (2008). Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS. Biol.* 6, e16.

Las, L., Stern, E. A., and Nelken, I. (2005). Representation of tone in fluctuating maskers in the ascending auditory system. *J. Neurosci.* 25, 1503–1513.

Masterton, R. B., Granger, E. M., and Glendenning, K. K. (1992). Psychoacoustical contribution of each lateral lemniscus. *Hear. Res.* 63, 57–70.

May, B. J., Huang, A., LePrell, G., and Hienz, R. D. (1996). Vowel formant frequency discrimination in cats: comparison of auditory nerve representations and psychophysical thresholds. *Aud. Neurosci.* 3, 135–162.

Moore, B. C. (2003). An Introduction to the Psychology of Hearing. London, Academic Press.

Moshitch, D., Las, L., Ulanovsky, N., Bar-Yosef, O., and Nelken, I. (2006). Responses of neurons in primary auditory cortex (A1) to pure tones in the halothane-anesthetized cat. *J. Neurophysiol.* 95, 3756–3769.

Nelken, I., and Ahissar, M. (2006). High-level and low-level processing in the auditory system: the role of primary auditory cortex. In Dynamics of Speech Production and Perception, P. Divenyi et al. eds (Amsterdam, IOS Press), pp. 343–353.

Nelken, I., Bizley, J. K., Nodal, F. R., Ahmed, B., Schnupp, J. W., and King, A. J. (2004). Large-scale organization of ferret auditory cortex revealed using continuous acquisition of intrinsic optical signals. *J. Neurophysiol.* 92, 2574–2588.

Nelken, I., and Chechik, G. (2007). Information theory in auditory research. *Hear. Res.* 229, 94–105.

Nelken, I., Rotman, Y., and Bar-Yosef, O. (1999). Responses of auditory-cortex neurons to structural features of natural sounds. *Nature* 397, 154–157.

Pelleg-Toiba, R., and Wollberg, Z. (1991). Discrimination of communication calls in the squirrel monkey: "call detectors" or "cell ensembles"? *J. Basic Clin. Physiol. Pharmacol.* 2, 257–272.

Phillips, D. P., and Orman, S. S. (1984). Responses of single neurons in posterior field of cat auditory cortex to tonal stimulation. *J. Neurophysiol.* 51, 147–163.

Qin, L., Kitama, T., Chimoto, S., Sakayori, S., and Sato, Y. (2003). Time course of tonal frequency-response-area of primary auditory cortex neurons in alert cats. *Neurosci. Res.* 46, 145–152.

Rauschecker, J. P., and Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc. Natl. Acad. Sci. USA* 97, 11800–11806.

Read, H. L., Winer, J. A., and Schreiner, C. E. (2002). Functional architecture of auditory cortex. *Curr. Opin. Neurobiol.* 12, 433–440.

Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., and Rauschecker, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136.

Schreiner, C. E., and Mendelson, J. R. (1990). Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J. Neurophysiol.* 64, 1442–1459.

Scott, S. K., and Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends Neurosci.* 26, 100–107.

Semple, M. N., and Kitzes, L. M. (1993). Focal selectivity for binaural sound pressure level in cat primary auditory cortex: two-way intensity network tuning. *J. Neurophysiol.* 69, 462–473.

Shackleton, T. M., Arnott, R. H., and Palmer, A. R. (2005). Sensitivity to interaural correlation of single neurons in the inferior colliculus of guinea pigs. *J. Assoc. Res. Otolaryngol.* 6, 244–259.

Shackleton, T. M., Skottun, B. C., Arnott, R. H., and Palmer, A. R. (2003) Interaural time difference discrimination thresholds for single neurons in the inferior colliculus of Guinea pigs. *J. Neurosci.* 23, 716–724.

Tan, A. Y., Zhang, L. I., Merzenich, M. M., and Schreiner, C. E. (2004). Tone-evoked excitatory and inhibitory synaptic conductances of primary auditory cortex neurons. *J. Neurophysiol.* 92, 630–643.

Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *J. Neurosci.* 24, 10440–10453.

Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat. Neurosci.* 6, 391–398.

Wang, D., and Brown, G. J. (2006). Fundamentals of computational auditory scene analysis. In Computational Auditory Scene Analysis: Principles, Algorithms and Applications, D. Wang and G. J. Brown eds (Hoboken, NJ, Wiley & Sons).

Wang, X., Lu, T., Snider, R. K., and Liang, L. (2005). Sustained firing in auditory cortex evoked by preferred stimuli. *Nature* 435, 341–346.

Wehr, M., and Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 426, 442–446.

Winkler, I., van Zuijen, T. L., Sussman, E., Horvath, J., and Naatanen, R. (2006). Object representation in the human auditory system. *Eur. J. Neurosci.* 24, 625–634.

Wollberg, Z., and Newman, J. D. (1972). Auditory cortex of squirrel monkey: response patterns of single cells to species-specific vocalizations. *Science* 175, 212–214.

Yost, W. A., and Sheft, S. (1993). Auditory perception. In Human Psychophysics, W. A. Yost et al. eds (New York, Springer), pp. 193–236.

Zatorre, R. J., Bouffard, M., Ahad, P., and Belin P. (2002) Where is 'where' in the human auditory cortex? *Nat. Neurosci.* 5, 905–909.