



# Music and the auditory brain: where is the connection?

Israel Nelken\*

Department of Neurobiology, The Interdisciplinary Center for Neural Computation and The Edmond and Lily Safra Center for Brain Research, Hebrew University, Jerusalem, Israel

**Edited by:**

Idan Segev, The Hebrew University of Jerusalem, Israel

**Reviewed by:**

Carles Escera, University of Barcelona, Spain

Robert J. Zatorre, McGill University, Canada

**\*Correspondence:**

Israel Nelken, Department of Neurobiology, Hebrew University, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel.  
e-mail: israel@cc.huji.ac.il

Sound processing by the auditory system is understood in unprecedented details, even compared with sensory coding in the visual system. Nevertheless, we do not understand yet the way in which some of the simplest perceptual properties of sounds are coded in neuronal activity. This poses serious difficulties for linking neuronal responses in the auditory system and music processing, since music operates on abstract representations of sounds. Paradoxically, although perceptual representations of sounds most probably occur high in auditory system or even beyond it, neuronal responses are strongly affected by the temporal organization of sound streams even in subcortical stations. Thus, to the extent that music is organized sound, it is the organization, rather than the sound, which is represented first in the auditory brain.

**Keywords:** music, auditory, brain, neurons

## MUSIC AND THE AUDITORY SYSTEM

When I started studying the auditory system, I claimed that I wanted to understand why monkeys prefer listening to the Rolling Stones rather than to Mozart. The monkeys I referred to were, obviously, not *Macaca mulatta*, but rather a subspecies of *Homo sapiens* that I disliked at the time. However, only today, many years later, I can perceive the real conceit of the younger me: the assumption that by studying the auditory system I will be able to understand the reactions of humans to the fine distinctions that separate Rock music from classical music. The “fine” should be taken seriously: a corollary of the argument I want to make in this perspective is that in terms of our current understanding of auditory processing, there is not much difference between the two.

Having already said “music,” “auditory processing,” and “understanding,” I have to define the scope of my argument. I will not try to define music beyond the trivial remark that while music has to do with sounds, not all sound is music. For example, I do not consider the “music of nature,” sounds in the natural environment, to be music, in the same sense that a magnificent sunset above the hills west of Jerusalem is not art. So what I am interested in has to do with the fact that someone took many sounds and organized them in some way – music includes not only sound, but also organization, both in sound space and in time.

Contrary to this vague definition of music, when I say “auditory processing,” I have something very definite in mind – I mean the biological processes (my bias being toward the electrical ones) that occur between the vibration of the tympanic membrane at one end and the spiking activity of neurons in the auditory brain at the other end. I will almost completely ignore here evidence from fMRI, which at best can give some hints as to the location of active neurons; and most evidence from EEG and MEG, which, while measuring electricity, reflect only distantly the actual spiking activity of neurons. Thus, my view of auditory processing in this perspective is unabashedly neuron-centric: by “understanding” I mean the reduction of (some of) the phenomenology of music

into neural mechanisms, spikes, synaptic currents, ion channels and all.

Finally, I have to define the parts of the brain I am considering as auditory. This is a surprisingly hard question. While the auditory nerve, the multiple brainstem auditory centers with their intricate analysis of auditory neural signals culminating in the hugely complex inferior colliculus, the medial geniculate body, and the primary auditory cortex are all clearly parts of the auditory brain, there are many other brain areas where there are auditory responses but which are not considered as auditory. These include, for example, the amygdala (Quirk et al., 1997), the superior colliculus (Middlebrooks and Knudsen, 1984), the hippocampus (Edeline et al., 1988), and the cerebellum (Huang and Burkard, 1986), to name just a few subcortical centers; and many cortical areas that lie beyond the classical auditory cortex (e.g., Cohen et al., 2004). For the purpose of this perspective, I will concentrate on the “core” auditory regions, those parts of the brain that would be considered as “the” auditory system in a textbook of the nervous system – the subcortical ascending auditory pathways, primary auditory cortex, and the surrounding fields.

I should immediately admit the limitations of this strongly reductionist approach. First, I am limiting myself to (mostly) data from animal studies. At the early stages of processing that I am considering, mammalian brains are reasonably similar to each other so that this is probably not a serious constraint. Second, a phenomenon as complex as music cannot be reduced to the responses of single neurons, but would require studying simultaneously the responses of many neurons distributed throughout the brain. Even an account as reductionist as the one I am considering here would require taking into account such brain-wide networks; however, my argument will be based on evidence from single-neuron responses only. Third, the brain areas I am considering are rarely those specifically activated by music in human imaging studies (e.g., Janata, 2005). As I will argue below, processing of relevance for music is performed in these areas,

in spite of the generally negative evidence from human imaging studies.

With these cautionary notes out of the way, here is the main argument of this perspective: with our current understanding of the auditory system, we stand in the paradoxical situation in which we do not understand “sound,” while we have a strong handle on “organization.” Thus, the low-level representations of sounds on which music is based are badly understood, and may in fact occur only “higher up” in the brain, outside the auditory brain as defined here. On the other hand, high-level aspects of music, such as sound organization in time, are strongly reflected within this same auditory brain.

## SOUNDS AND THE AUDITORY BRAIN

I will pursue the first part of my argument in two ways. I will first argue that we do not quite understand where and how the low-level properties of sounds, such as pitch and timbre, are represented in the auditory system. I will then argue that this is really a side effect of an even larger gap in our understanding – the fact that we do not understand the relationships between the pressure waves that cause the tympanic membrane to vibrate, and the introspective percept we call sound, which is very far removed from both the physical vibration that initiated it and from the representation of these vibrations in the auditory system (at least as defined here).

Let us consider sound processing in the auditory system and its relationships to a fundamental property of sound that is used in music – pitch. Pitch is without doubt one of the most important properties of sounds with which humans do music. The major physical correlate of pitch is periodicity (not frequency!), but this is not an absolute identification – there are periodic sounds that do not elicit pitch at their periodicity, and non-periodic sounds that do elicit pitch (Schnupp et al., 2011, Chapter 3). Most importantly, pitch represents an abstraction: many different sounds have the same pitch (e.g., a violin, cello, trumpet, flute, and a piano all playing the same note, see <https://mustelid.physiol.ox.ac.uk/drupal/?q=topics/same-melody-different-timbre>).

This abstract quality of pitch has consequences to our understanding of the coding of pitch in the auditory brain. To start with, it is often argued that since auditory nerve fibers follow the periodicity of sounds evoking pitch, pitch is coded in the auditory nerve. I believe that this is seriously wrong.

The heart of the matter is the fact that periodicity may depend on spectral content in a wide frequency band, while auditory nerve fibers are narrowly tuned; in general a single auditory nerve fiber simply does not “hear” enough of the sound in order to respond to the right periodicity. Thus, a neuron whose best frequency is 200 Hz will respond roughly similarly to a sound with a periodicity of 100 Hz containing a prominent second harmonic and to a sound with a periodicity of 200 Hz with a prominent fundamental, and may not respond at all to a sound with a pitch of 200 Hz which misses its first few harmonics. In other words, the response of an auditory nerve fiber tuned to 200 Hz is neither sufficient nor necessary for a sound to be perceived as having a pitch of 200 Hz.

This fact is well known, but is usually handled by claiming that it is the activity in the whole array of auditory nerve fibers that represents the pitch of a sound. This claim is in a way true – by observing the array of auditory nerve fibers, it should certainly

be possible to determine the pitch of a sound. After all, when we listen to sounds, we extract pitch from the auditory nerve activity pattern all the time. However, this claim also misses the point, in two ways.

First, such a claim does not solve the problem of the coding of pitch. Somewhere in the brain, some structure has to take the array of activity of the auditory nerve fibers, and use it to extract the invariant representation of pitch (or so we intuit), so claims about “population coding” just shove the problem of pitch coding away without solving it. There is no extra explanatory power in the claim that the auditory nerve fibers represent pitch relative to the claim that the pressure vibrations in the air represent pitch.

Second, and possibly even more importantly, the array of auditory nerve fibers represents not only pitch, but also all other perceptual properties of sounds. The same fibers whose responses contain information about the pitch also carry information about timbre and loudness. In fact, in as much as we can talk about representations in the auditory nerve, the array of auditory nerve fiber represents very clearly one thing – the physical vibrations of the tympanic membrane. It does not even represent the abstract quantity called periodicity, not to mention the perceptual quality called pitch.

What about stations higher up in the auditory pathway? There is a substantial and important work on the coding of pitch in the brainstem. As in the case of auditory nerve fibers, brainstem neurons follow the periodicity of the acoustic stimulus, but the dominant sound representations all the way up to the inferior colliculus share with the auditory nerve fibers the narrow width of tuning of each individual element and the high sensitivity to many (if not all) properties of sounds. Thus, while periodic sounds evoke strong periodic activity in the brainstem (e.g., Winter et al., 2001), there is no convincing evidence that the brainstem (even the inferior colliculus) has an explicit representation of pitch (Reviewed in Schnupp et al., 2011, Chapter 3). In fact, the most convincing suggestions for the structure(s) that perform this abstraction, from sounds to their pitch, are far up in the auditory hierarchy, certainly above primary auditory cortex, both in humans (Patterson et al., 2002; Hall and Plack, 2009) and in non-human primates (Bendor and Wang, 2005). This is, in a way, a rather surprising finding. Sounds go all the way from the periphery to primary auditory cortex and above without an explicit assignment of pitch. And without a pitch representation, it is hard to see how music can be represented.

I believe, however, that the gap between music and the current understanding of the auditory system is much wider than this upside-down result. In my discussion of pitch coding, I ignored a crucial facet of real-world sounds: contrary to most auditory experiments (including many of mine), we usually hear more than one “sound” at each moment in time. For example, while typing this manuscript, I hear the low rumble of the power supplies of the many computers in my lab, a merle singing outside the window of my office, and the tick tack of the keys I hit while typing. My auditory nerve fibers carry information about the mixture, not about any particular component of it. There is an important corollary here – at the level of the auditory nerve, many pitches may be present concurrently. In as much as this is music, these different pitches have at least some individual existence. However, it is hard

to think of ways of identifying the different pitches without at the same time also separating out the different bits of sounds that sum up to produce the mixture at the ear (Schnupp et al., 2011, Chapter 6).

This argument puts in the foreground the need to understand the transformation that occurs in the brain between the physical stimuli and the “objects of perception,” those things that carry the perceptual properties we attribute to sounds such as pitch, timbre, spatial location, and so on. Music is done to a significant degree with these “objects of perception” – the individual tones composing a chord, melody as separate from its accompaniment, and so on and so forth. In the last 10 years or so, electrophysiologists studying the auditory brain came to appreciate the great importance of a loose collection of processing tasks called auditory scene analysis (ASA) whose goal is to form these objects of perception (Bregman, 1990). In fact, I consider ASA, in a wide enough sense, as the major processing task of the auditory system. Thus, understanding how neurons do ASA is a necessary step toward understanding music in the brain.

So, how much do we understand ASA in neuronal terms? Consider one important advance in understanding the implementation of ASA in the brain: the recent spate of work on streaming. In a typical streaming experiment, two sounds are presented alternately to the subject. If the difference between the two sounds (e.g., frequency separation between two pure tones) is large enough, and/or if they are played fast enough, the sequence of sounds breaks down perceptually into two “things,” each containing one of the two sounds (hear the illustration at <https://mustelid.physiol.ox.ac.uk/drupal/?q=topics/streaming-galloping-rhythm-paradigm>). Bregman (1990) named the two “things” streams. The groundbreaking work of Fishman et al. (2001) led to a neural account for the breakdown process of the single sequence of pure tones into two streams: they showed that under conditions in which a single sequence is heard, neurons in auditory cortex of macaques tend to respond to both tones, while when a breakdown occurs they tend to respond to either one tone or the other. Using these ideas, Micheyl et al. (2005) showed that the dynamics of the breakdown process in human listeners can be accounted for by the dynamics of neural responses in auditory cortex of macaques. Recently, Elhilali et al. (2009) remarked that there should be also an important role for the temporal incoherence of the neuronal responses to the two tones in the two-stream condition, adding yet another component to the neural model of streaming.

While these are significant advances in the process of linking the perceptual phenomenon of streaming with neural responses, it is important to realize that these studies did not find a neural representation of streams. The neurons studied by Fishman, Micheyl, Elhilali and their colleagues just responded to the individual sounds in the sequence. Instead, these studies demonstrate properties of neural responses that may be used by a hypothetical (but at this point possibly mythical) next layer to create streams. Thus, important as they are, these studies do not solve the issue of the representation of streams in the auditory system.

To the best of my knowledge, there is only one non-trivial example of the end-product of ASA in neural hardware: the specific responses of neurons in cat auditory cortex to the background components of natural sounds (Bar-Yosef and Nelken,

2007; Nelken and Bar-Yosef, 2008). In these experiments, short segments of natural recordings of bird songs have been used. These segments were digitally processed to remove the bird songs, preserving only the background rustling. Many neurons responded to the natural sound with similar responses to those they emitted when presented with the background alone, but responded differently when presented with the clean bird song. These neurons respond to one bit of the sound independently of the presence of other bits of sounds, which may be substantially louder inside their frequency response area. Unfortunately, the neural mechanisms leading to such responses have not been worked out.

## ORGANIZATION AND THE AUDITORY BRAIN

Here comes what is, for me, the most surprising twist in the plot. There is in fact significant amount of processing in the auditory brain which I find highly relevant for music. However, it does not have much to do with the “sound” of music, but rather with the “organization” of music.

The phenomena I want to emphasize here occur at a time scales of seconds to minutes. Responses of neurons to sounds turn out to depend on the recent history of sound presentations. Early clues to these effects have been known for many years. For example, Condon and Weinberger (1991) showed a strong depression in the response to a frequency that pipped continuously for a few minutes, but this depression did not generalize to other, nearby tone frequencies, and therefore did not represent a “fatigue” of the neuronal responses.

It was however the introduction of the oddball paradigm into single-neuron studies by Ulanovsky et al. (2003) that really energized the study of context sensitivity in the auditory system. The oddball paradigm has been extensively used in human studies (Naatanen et al., 2007) to study the important component of the auditory event-related potentials called mismatch negativity (MMN). Ulanovsky et al. (2003) adapted this paradigm to single-neuron studies. In a typical oddball experiment, two tones are presented in a sequence, one common and one rare. In a different sequence, the two tones are again presented but with their roles reversed. The typical result of such experiments is that the response to the same tone may be substantially larger when rare than when common. This effect, named “stimulus-specific adaptation” (SSA, Ulanovsky et al., 2003) when considered in the context of single-neuron responses, has been now studied by a large number of groups and shown to be present in auditory cortex of anesthetized cats (Ulanovsky et al., 2003); awake rats (Von Der Behrens et al., 2009); the inferior colliculus of rats (Malmierca et al., 2009); and the medial geniculate body of rats and mice (Anderson et al., 2009; Antunes et al., 2010).

Stimulus-specific adaptation is relevant to music because it shows that the responses of neurons to the same sound depend on the organization of its recent past. In the case of oddball sequences, this is a rather simple dependency – the less common the sound, the larger the response it evokes. However, recent work in my laboratory (Taaseh et al., 2011) compared the responses evoked by the same tones in a number of different sequences, showing for example that the responses to a rare tone played with a common sound are shaped by different mechanisms than the responses to a tone that is rare, but played together with many different sounds,

all of whom are rare. Similarly, the responses to the two tones in an oddball sequence depend on whether the sequence is regular, with fixed intervals between presentations of the rare tone, or whether the sequence is random, with a fixed probability of presentation of the rare tone. Thus, the neuronal responses in auditory cortex do not depend only on the probability of the tones, but also on fine details of their order (Nelken et al., 2010). Furthermore, SSA is not limited to pure tones – frozen tokens of white noise evoke SSA when played in an oddball configuration (Nelken et al., 2010). Thus, SSA in primary auditory cortex of anesthetized rats seems to engage mechanisms that are sensitive to the detailed history of the sound sequence, and not only to the rarity of the rare pure frequency tone.

## SO WHAT?

The auditory brain shows little evidence of sound representations in terms of their perceptual qualities, and, even worse, does not even seem to represent sounds at all, at least in the usual everyday sense of the use of the word “sound.” Instead, the auditory brain seems to represent, quite well, the physical vibrations of the tympanic membrane. At the level of the auditory cortex there are some hints of representations that either emphasize features of sounds that can be used later to create “objects of perception” or “streams” (I am vague on purpose), or even already separates

sound mixtures into their components. It is in this sense that classical and Rock music are not that different from each other. Most low-level descriptors of the two would not be too far from each other – overall spectral range (with a possible advantage to Rock music at very low frequencies), typical rates of spectral and temporal modulations, and all of these other properties that modulate the responses of auditory neurons in the early parts of the auditory brain.

While we struggle with the nature of sound representations in the auditory brain, it is singularly easy to observe the signature of sound organization on the neural responses, starting as early as the inferior colliculus. Thus, organization is reflected in the neural responses of the auditory brain more strongly, and at earlier stages, than sounds (in the sense of “objects of perception”). This is non-intuitive (at least to me). Taken to the extreme, this state of affairs may mean that in the “organized sound” that music may be, we may have easier time accounting for the “organized” than for the “sound” within the confines of the auditory brain. Thus, it may well be that when brains process music, organization comes first, and sound only follows.

## ACKNOWLEDGMENTS

This work was supported in part of grants from the Israeli Science Foundation (ISF) and the German-Israeli Foundation (GIF).

## REFERENCES

- Anderson, L. A., Christianson, G. B., and Linden, J. F. (2009). Stimulus-specific adaptation occurs in the auditory thalamus. *J. Neurosci.* 29, 7359–7363.
- Antunes, F. M., Nelken, I., Covey, E., and Malmierca, M. S. (2010). Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *PLoS ONE* 5, e14071. doi:10.1371/journal.pone.0014071
- 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.* 1:3. doi:10.3389/neuro.10.003.2007
- Bendor, D., and Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature* 436, 1161–1165.
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.
- Cohen, Y. E., Cohen, I. S., and Gifford, G. W. III. (2004). Modulation of LIP activity by predictive auditory and visual cues. *Cereb. Cortex* 14, 1287–1301.
- Condon, C. D., and Weinberger, N. M. (1991). Habituation produces frequency-specific plasticity of receptive fields in the auditory cortex. *Behav. Neurosci.* 105, 416–430.
- Edeline, J. M., Dutrieux, G., and Neuenschwander-El Massioui, N. (1988). Multiunit changes in hippocampus and medial geniculate body in free-behaving rats during acquisition and retention of a conditioned response to a tone. *Behav. Neural Biol.* 50, 61–79.
- Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J., and Shamma, S. A. (2009). Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron* 61, 317–329.
- Fishman, Y. I., Reser, D. H., Arezzo, J. C., and Steinschneider, M. (2001). Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hear. Res.* 151, 167–187.
- Hall, D. A., and Plack, C. J. (2009). Pitch processing sites in the human auditory brain. *Cereb. Cortex* 19, 576–585.
- Huang, C. M., and Burkard, R. (1986). Frequency sensitivities of auditory neurons in the cerebellum of the cat. *Brain Res.* 371, 101–108.
- Janata, P. (2005). Brain networks that track musical structure. *Ann. N. Y. Acad. Sci.* 1060, 111–124.
- Malmierca, M. S., Cristaudo, S., Perez-Gonzalez, D., and Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J. Neurosci.* 29, 5483–5493.
- Micheyl, C., Tian, B., Carlyon, R. P., and Rauschecker, J. P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron* 48, 139–148.
- Middlebrooks, J. C., and Knudsen, E. I. (1984). A neural code for auditory space in the cat's superior colliculus. *J. Neurosci.* 4, 2621–2634.
- Naatanen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590.
- Nelken, I., and Bar-Yosef, O. (2008). Neurons and objects: the case of auditory cortex. *Front. Neurosci.* 2:1. doi:10.3389/neuro.01.009.2008
- Nelken, I., Yaron, A., and Taaseh, N. I. H. (2010). “Responses to sound sequences in auditory cortex: beyond the oddball paradigm,” in *40th Annual Meeting of the Society for Neuroscience*, San Diego, CA.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., and Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36, 767–776.
- Quirk, G. J., Armony, J. L., and Ledoux, J. E. (1997). Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. *Neuron* 19, 613–624.
- Schnupp, J., Nelken, I., and King, A. J. (2011). *Auditory Neuroscience: Making Sense of Sound*. Cambridge, MA: MIT Press.
- Taaseh, N., Yaron, A., and Nelken, I. (2011). Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS ONE* 6: e23369. doi: 10.1371/journal.pone.0023369
- Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat. Neurosci.* 6, 391–398.
- Von Der Behrens, W., Bauerle, P., Kossel, M., and Gaese, B. H. (2009). Correlating stimulus-specific adaptation of cortical neurons and local field potentials in the awake rat. *J. Neurosci.* 29, 13837–13849.
- Winter, I. M., Wiegand, L., and Patterson, R. D. (2001). The temporal representation of the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig. *J. Physiol. (Lond.)* 537, 553–566.

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

Received: 19 May 2011; accepted: 10 September 2011; published online: 27 September 2011.

Citation: Nelken I (2011) Music and the auditory brain: where is the connection? *Front. Hum. Neurosci.* 5:106. doi: 10.3389/fnhum.2011.00106

Copyright © 2011 Nelken. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.