Neural encoding of sensory and behavioral complexity in the auditory cortex

Kishore Kuchibhotla[1,2] and Brice Bathellier[3]

Converging evidence now supports the idea that auditory cortex is an important step for the emergence of auditory percepts. Recent studies have extended the list of complex, nonlinear sound features coded by cortical neurons. Moreover, we are beginning to uncover general properties of cortical representations, such as invariance and discreteness, which reflect the structure of auditory perception. Complexity, however, emerges not only through nonlinear shaping of auditory information into perceptual bricks. Behavioral context and task-related information strongly influence cortical encoding of sounds via ascending neuromodulation and descending top-down frontal control. These effects appear to be mediated through local inhibitory networks. Thus, auditory cortex can be seen as a hub linking structured sensory representations with behavioral variables.

Addresses
[1] Department of Psychological and Brain Sciences, Department of Neuroscience, Johns Hopkins University, Baltimore, MD 21218, United States
[2] Laboratoire de Neurosciences Cognitives, INSERM U960, École Normale Supérieure – PSL Research University, Paris, France
[3] Unité de Neuroscience, Information et Complexité (UNIC), FRE 3693, Centre National de la Recherche Scientifique and Paris-Saclay University, Gif-sur-Yvette, 91198, France

Corresponding author: Bathellier, Brice (brice.bathellier@unic.cnrs-gif.fr)

Introduction

Auditory perception is an essential, complex process through which humans and other animals use pressure waves to interpret and interact with the environment. The cochlea initially decomposes pressure waves into their frequency spectrum and transduces these signals into neural impulses [1]. This tonotopic organization propagates throughout the auditory system. At the level of perception, however, audition is not merely designed for precise frequency encoding [2], but rather interprets complex acoustic motifs into distinct auditory objects reflecting our experience of the acoustic environment. Thus, one may struggle to identify the absolute pitch of a tone but most people accustomed to western music will recognize if the tone is coming from a piano [3]. How does the auditory system turn the initial spectral decomposition of sounds into the timbre of instruments, the words of a phrase, or the siren of a police car? In vision science, Gestalt theories [4,5] argue that perception is based on a skeleton of complex representations which does not correspond to a 'one-to-one copy' of the input signals but rather to archetypal building blocks that are innate or experience-dependent and are used to construct perception [2,6]. This idea likely extends to audition [7–9]. A major challenge for understanding auditory perception is to experimentally isolate and mechanistically explain the elements of such complex auditory representations. These sensory percepts, moreover, are bound to be sensitive to the current internal state and near-term goals of the subject while it explores or interacts with the environment. Increasing evidence now shows that, in auditory cortex, an elaborate vocabulary of auditory representations comes along with information about behavioral context. The purpose of this review is to highlight recent discoveries about these two levels of complexity in auditory cortex and discuss their functional roles and potential interactions.

Structure of auditory cortex representations and their link to perception

Whether auditory cortex represents simple acoustic features or building blocks for 'auditory objects' is a long-standing question [7,10–13]. The existence of frequency tuned neurons and their tonotopic mapping in primary auditory subfields [14,15] suggests a degree of similarity to cochlear representations. Yet this might be the only similarity. The precision of the tonotopic map weakens in secondary areas [15]. Moreover, frequency selectivity is, in general, more broadly tuned in auditory cortex than in subcortical areas [16,17], and response time constants appear to be longer [7] so that fine temporal details present in the input are lost or converted into rate codes [18•]. This readily suggests that cortex integrates over a wider range of information than subcortical structures. But most importantly, multiple studies have shown that selectivity of auditory cortex neurons greatly extend pure tone frequency coding [13,19–22]. Recently, two-photon calcium imaging and electrophysiology in rodents demonstrate that loosely mapped ensembles of cortical neurons in primary and secondary auditory cortex code for the

Figure 1
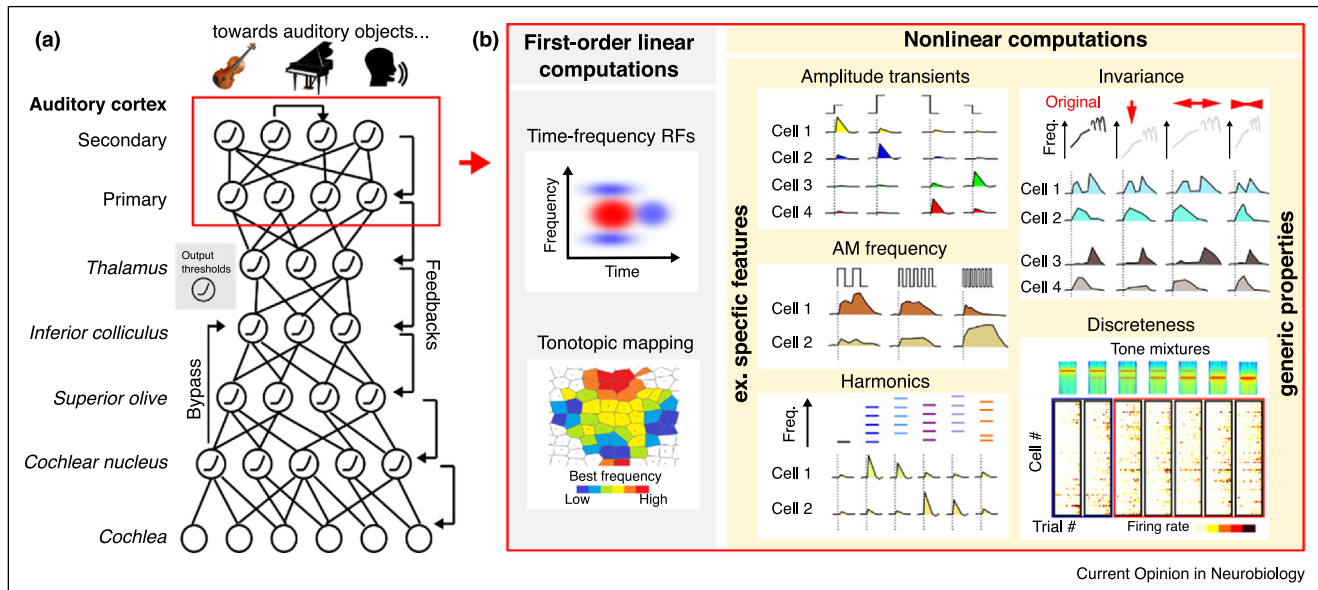
Structuring of auditory information via nonlinear computations in the auditory cortex. **(a)** Throughout the auditory system raw cochlear inputs are structured into biologically relevant percepts (auditory objects). This transformation requires a complex ensemble of non-linear computations, which we here schematize as a multilayer network linking simple nonlinearities (e.g. spike threshold) with an elaborate connection graph that includes feedforward, feedback and lateral connections. Interestingly, appropriately trained multilayer networks (Deep Learning) were recently demonstrated to boost performance on artificial perceptual tasks such as speech recognition. **(b)** Beyond simple computations such as spatially organized preference for particular frequency ranges (*left*), the auditory cortex displays a number of non-linear computations. These leads to the emergence of neurons sensitive to specific features, as sketched in the middle column, including sound onsets and offsets of particular amplitude [26••], amplitude modulation (AM) frequencies [25], or even harmonicity (at least in primates) [30••]. Also, generic response properties, such as invariance to modification of basic acoustic parameters as sketched in the *top right* graph [31] and discrete coordinated population response switches (*bottom right*) (data from [33]), indicate that these nonlinear computations endow cortical representations with some of the expected properties of auditory object representations.

direction of frequency variations [23], interaural differences [24•] as well as the frequency [25], amplitude and direction [26••] of intensity variations (Figure 1). These temporal features are crucial for recognizing particular classes of sounds. For example, recognition of musical instruments greatly depends on the steepness of tone intensity rise and decay [3] which is coded in primary auditory cortex, even in mice [26••]. Frequency modulations are important components of vocalization in most species [27–29]. For highly vocal animals, the structure of frequency harmonics represent important cues, as in human language [12]. Interestingly, a recent study showed that the core auditory cortex of primates includes cells which specifically detect patterns of frequency harmonics [30••] (Figure 1).

The presence of complex features in auditory cortex, however, does not prove that it codes for auditory objects. To address this issue, two major and somewhat orthogonal properties of object-like representations must be observed: invariance and discreteness. Invariance refers to the stability of representations with respect to small changes in acoustic parameters. Discreteness refers to the categorical and rapid switching of representations

between objects as separate entities. Two recent studies in rats have shown that auditory cortex neurons respond to vocalizations or water sounds with a certain degree of robustness against various acoustic modifications [31,32]. Interestingly, invariance for vocalization was tested both in primary and non-primary auditory cortex and was found to be more pronounced in non-primary areas [31] (Figure 1), suggesting that invariance properties progressively emerge along the cortical hierarchy, correlating with the weakening of the tonotopic map [15]. As for discreteness, a two-photon calcium imaging study has shown that local ensemble of neurons in the mouse auditory cortex respond in a step-wise manner to gradual changes in sound mixtures [33] (Figure 1). The object-like representation in cortex predicted how mice categorized diverse sounds during a behavioral task [33]. Thus, many of the ingredients necessary to build object-like or categorical representations are present in auditory cortex and evidence exists that cortical representations are close to perceptual space.

Are object-like representations hard-wired into the cortex or are they experience-dependent, dynamically updating based on the ecological needs of a particular animal? In