

Perceptual learning in the developing auditory cortex

Shaowen Bao

Department of Physiology, University of Arizona, Tucson, AZ 85724, USA

Keywords: auditory cortex, categorical perception, development, perceptual learning, statistical learning

Abstract

A hallmark of the developing auditory cortex is the heightened plasticity in the critical period, during which acoustic inputs can indelibly alter cortical function. However, not all sounds in the natural acoustic environment are ethologically relevant. How does the auditory system resolve relevant sounds from the acoustic environment in such an early developmental stage when most associative learning mechanisms are not yet fully functional? What can the auditory system learn from one of the most important classes of sounds, animal vocalizations? How does naturalistic acoustic experience shape cortical sound representation and perception? To answer these questions, we need to consider an unusual strategy, statistical learning, where what the system needs to learn is embedded in the sensory input. Here, I will review recent findings on how certain statistical structures of natural animal vocalizations shape auditory cortical acoustic representations, and how cortical plasticity may underlie learned categorical sound perception. These results will be discussed in the context of human speech perception.

Introduction

Learning has been defined as an enduring change in the mechanisms of behavior that results from experience with the environmental events (Domjan, 2010). Perceptual learning is the specific and relatively permanent modification of perception and behavior following sensory experience (Fahle & Poggio, 2002). Exposure to specific acoustic experience in the critical period of early sensory development alters cortical sound representations (Zhang *et al.*, 2001) and perceptual behavior (Han *et al.*, 2007; Kover *et al.*, 2013), and therefore is a process of perceptual learning. However, this type of perceptual learning through sensory exposure is unique in that it does not involve an explicit training process (Keuroghlian & Knudsen, 2007); there is no instruction of the desired response or feedback on the actual response. In the absence of instructions or feedbacks, how does the auditory system know what and how to learn in order to adapt to its specific acoustic environment?

Perceptual learning may be classified into three categories: unsupervised, supervised and reinforcement learning. In reinforcement learning, feedback is provided on whether the response is correct (e.g., tell me what sound you are hearing; no, that is not correct). In supervised learning, the desired response is provided (e.g., now I am going to play the sound of /la/; learn it.). Reinforcement learning is sometimes regarded as a form of supervised learning. In unsupervised learning, however, the subject is left alone to discover the structures of a stimulus ensemble without instructions or feedback (e.g., I am going to play a long sound stream; tell me all the structures in it that you can find). This is achieved through statistical learning, a process of reshaping perception according to the statisti-

cal structures of a stimulus ensemble (e.g., an acoustic environment). Although statistical learning *per se* does not need feedback, it can be performed in a reinforcement context (Toro & Trobalon, 2005).

Both humans and rodents are sensitive to statistical structure of acoustic input such as stimulus probability and conditional probability distributions. For example, exposure to sounds distributed along a phonetic continuum affects the subsequent discrimination of those phonemes in rats and humans (Maye *et al.*, 2002; Pons, 2006). Human infants and adults are sensitive to stimulus transitional probability, and can use it to identify speech sound sequences or tone sequences that are repeating in a continuous acoustic stream (Saffran *et al.*, 1996, 1999). Rats are also sensitive to conditional probability such as co-occurrence of sounds in a sequence (Toro & Trobalon, 2005). The neural mechanisms underlying this type of statistical learning are unknown.

The auditory cortex is remarkably adaptive to sensory input. During an epoch of early development, exposure to the acoustic environment can change sound representations without external instructions or feedback (Zhang *et al.*, 2001). Early studies indicate that cortical sound representations are sensitive to simple statistics of the sensory input such as frequency of occurrences: more frequently experienced sounds gain larger cortical representations (Zhang *et al.*, 2001). More recent studies suggest that auditory cortex is also sensitive to conditional probabilities (e.g., co-occurrence of sounds in a sequence; Kover *et al.*, 2013). Most importantly, early experience-dependent reorganization of cortical acoustic representations is correlated with altered perception and perceptual behavior (Han *et al.*, 2007; Kover *et al.*, 2013). Here I will discuss recent findings on the cortical mechanisms underlying developmental perceptual learning. For more comprehensive reviews on cortical plasticity and its perceptual consequences, please see Sanes & Bao (2009) and Schreiner & Polley (2014).

Correspondence: Shaowen Bao, as above.
E-mail: sbao@email.arizona.edu

Received 29 September 2014, revised 30 November 2014, accepted 6 December 2014

The developing auditory cortex selectively represents animal vocalizations

Nature sounds typically comprise environmental sounds (e.g., wind, water), animal vocalizations (including human speech) and non-vocalization animal sounds (e.g., from footsteps, wing flaps). These sounds carry different behavioral significances. Animal vocalizations of the same and different species are likely to be crucial for reproduction and survival of the animal. By contrast, environmental sounds are likely to be less important. Human infants preferentially attend to speech over non-speech sounds (Vouloumanos & Werker, 2007), suggesting an intrinsic preference for behaviorally important sounds. At the level of sensory neural processing, it would be advantageous to preferentially allocate more neurons to process behaviorally important sounds such as animal vocalizations (Wang & Kadia, 2001; Garcia-Lazaro *et al.*, 2006; Kim & Bao, 2013).

Among the different types of natural sounds, animal vocalizations are arguably the most structured (Singh & Theunissen, 2003). They are complex and diverse, but also have some common characteristics that distinguish them from non-vocalization sounds. For example, most vocalizations of mammals are repeated at an ethological range of 4–10 Hz (Liu *et al.*, 2003; Schnupp *et al.*, 2006). Human speech is also temporally modulated (Rosen, 1992), and the temporal modula-

tion in the same ethological range is critical for human speech perception (Elliott & Theunissen, 2009). By contrast, non-vocalization sounds are often random in nature and are not repeated in the ethological modulation range. Studies have indicated that sounds that are repeatedly encountered in the ethological rate range become over-represented in the auditory cortex, i.e. more neurons become tuned to the sounds (Zhang *et al.*, 2001; Chang & Merzenich, 2003; de Villers-Sidani *et al.*, 2008; Zhou *et al.*, 2008; Insanally *et al.*, 2009). Sounds that are repeated at higher or lower rates are not over-represented (Fig. 1 and Kim & Bao, 2009). Indeed, unmodulated sounds, similar to a constant environmental sound, can retard cortical development (Chang & Merzenich, 2003; de Villers-Sidani *et al.*, 2008; Zhou *et al.*, 2008). Unmodulated sounds may even be under-represented in the developing auditory cortex (de Villers-Sidani *et al.*, 2008; Zhou *et al.*, 2008). This type of temporal repetition rate-dependent cortical plasticity probably contributes to over-representation of conspecific vocalizations in rat auditory cortex (Fig. 1 and Kim & Bao, 2013).

Statistics of the acoustic input shape cortical sound representations

Humans and rodents are sensitive to stimulus statistics such as simple stimulus probability distributions and more complex transitional

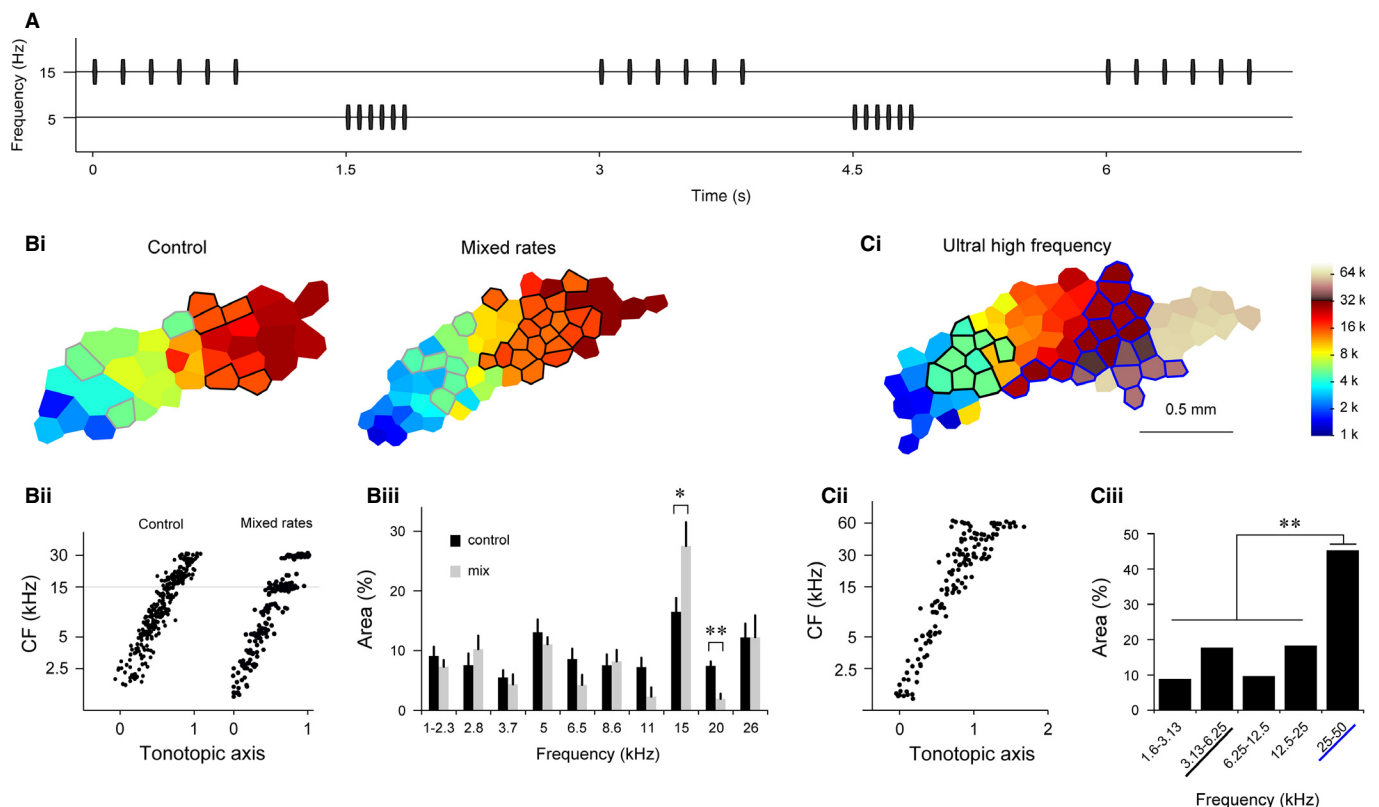


FIG. 1. Over-representation of sounds repeated at the ethological rates. (A) A schematic of 'mix-rate' rearing stimuli. A train of 15-kHz tones consisted of six tone pips presented at the ethological rate (6 Hz), and a train of 5-kHz tones consisted of six tone pips presented at the fast rate (15 Hz). Trains of the two repetition rates were interleaved such that one train was heard every 1.5 s. (B) Characteristic frequency (CF) map reorganization resulted from the mixed-rate rearing. (Bi) Example maps of control and mixed-rate animals. Control animal is the same as seen in Fig. 2A. Areas representing 5 kHz \pm 0.2 octaves are outlined in gray while areas representing 15 kHz \pm 0.2 octaves are outlined in black. (Bii) Distributions of CFs along the tonotopic axis. (Biii) Sizes of cortical areas representing different frequency bands. There was a significant increase in representation at 15 kHz and a significant decrease at 20 kHz. (C) Cortical representation of ultrasonic frequencies. (Ci) An example CF map from a control animal mapped up to 74 kHz. Areas representing 25–50 kHz are outlined in blue while areas representing 3.13–6.25 kHz are outlined in black. (Cii) Distribution of CFs along the tonotopic axis. (Ciii) Sizes of cortical areas representing one-octave frequency bands. The representation of the 25–50 kHz band was significantly larger than those of the other. Error bars depict SEM. * $P < 0.05$, ** $P < 0.001$. This figure was originally published in Kim & Bao (2009).

or conditional probability distributions (Saffran *et al.*, 1996, 1999; Maye *et al.*, 2002; Toro & Trobalon, 2005; Pons, 2006). Electrophysiological studies suggest that developing auditory cortex can utilize these statistics to shape acoustic representations. Early studies indicate that more neurons become tuned to repeatedly presented (i.e., high probability) sounds (Zhang *et al.*, 2001; Chang & Merzenich, 2003; de Villers-Sidani *et al.*, 2008; Zhou *et al.*, 2008; Insanally *et al.*, 2009). It may be hypothesized that the size of cortical representation encodes stimulus probability (Simoncelli, 2009; Fischer, 2010; Kover & Bao, 2010). A recent study indicated that developing auditory cortex can also encode higher-order conditional stimulus probability, i.e. the probability of sounds occurring in a sequence played at the ethological rate (Kover *et al.*, 2013). Sounds that are presented in a sequence tend to be represented by the same population of neurons, whereas sounds that are never presented in a sequence tend to be represented by separate populations of neurons (Fig. 2 and Kover *et al.*, 2013). Rodent vocalizations occur in bouts with similar temporal rates (Liu *et al.*, 2003; Holy & Guo, 2005; Kim & Bao, 2009, 2013). Early experience of those vocalizations

could thus theoretically lead to similar cortical representations, and reduced perceptual contrast, of the individual calls despite their substantial trial-by-trial variability. By contrast, functionally different call types (e.g., pup vs. adult encounter calls) that do not occur in the same bout may be represented by distinct populations of neurons, resulting in perceptual boundaries and categorical perception of the calls (Ehret & Haack, 1981).

Experience-dependent cortical plasticity can account for altered perceptual behaviors

How does statistical learning in the developing auditory cortex impact perception and perceptual behavior? Early acoustic experience has a profound impact on auditory perception and perceptual behaviors. Human fetuses gradually become sensitive to mother's voice and native speech during late gestation (Kisilevsky *et al.*, 2009; Kisilevsky & Hains, 2011), presumably due to acoustic experience *in utero*. Prenatal experience shapes perception and neural responses to speech in neonatal infants (Nazzi *et al.*, 1998; Partanen

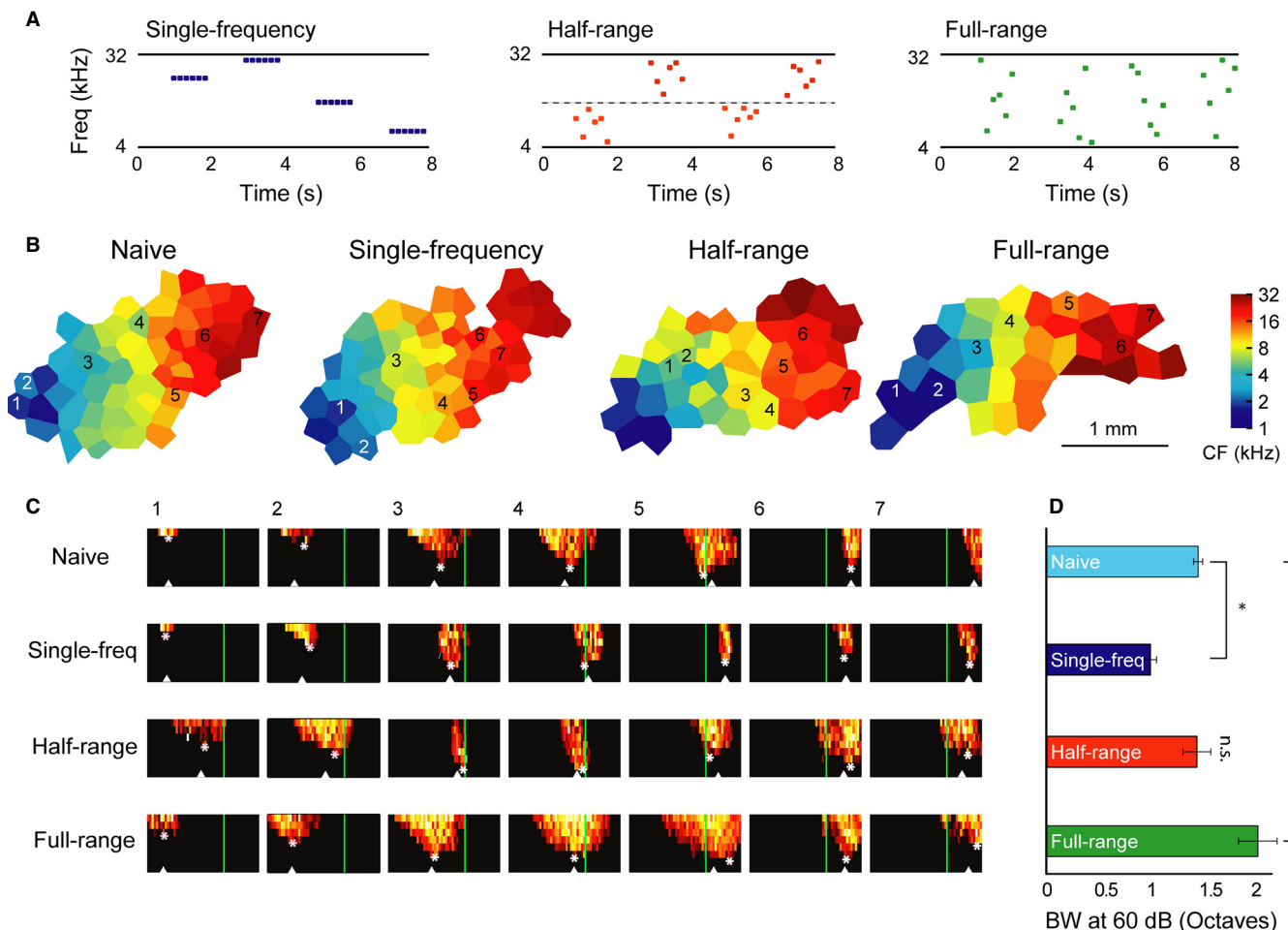


FIG. 2. Influences of higher-order stimulus statistics on spectral selectivity of primary auditory cortical neurons. (A) Schematics of the acoustic environments that the animals experienced. The three acoustic environments had the same logarithmically uniform frequency distribution from 4 to 32 kHz and the same temporal presentation rates, but differed in the conditional probabilities of the tonal frequencies within sequences. (B) Representative cortical maps. The sound exposure did not alter the overall tonotopic characteristic frequency distribution. (C) Representative frequency-intensity receptive fields. The corresponding locations are marked on the tonotopic maps in B. The green vertical lines mark the low conditional probability boundary experienced by the half-range group. Stars denote the characteristic frequency (CF) and triangles denote the center-of-mass frequency. Horizontal axis depicts frequency logarithmically from 1 to 32 kHz and vertical axis depicts intensity from 10 to 80 dB SPL. (D) Tuning bandwidth at 60 dB SPL. Cyan, naive control; dark blue, single-frequency; red, half-range; green, full-range. Frequency tuning bandwidth became narrower in the single-frequency group and broader for the full-range group compared to control. Error bars depict SEM. * $P < 0.05$ determined by an ANOVA with *post hoc* Tukey–Kramer test. This figure was originally published in Kover *et al.* (2013).

et al., 2013). Language-specific perception and neural representation of speech sounds continue to refine and consolidate during the first year of life and beyond (Kuhl *et al.*, 1992, 2006). A profound consequence of early experience of speech sound is the sharpening of categorical perception of native speech sounds: within-category perceptual contrast is reduced and between-category contrast is enhanced, resulting in more efficient recognition of native speech sounds and often loss of sensitivity to some foreign phonemic contrasts (Kuhl *et al.*, 1992, 2006; Iverson *et al.*, 2003). Can cortical statistical learning as shown in electrophysiological studies result in categorical perception of conspecific vocalizations (Ehret & Haack, 1981)?

An early study examined perceptual consequences of early acoustic experience in rats (Han *et al.*, 2007). Rats were exposed to a 7.1-kHz tone repeated at the ethological rate during the critical period of auditory cortical development. The animals were then placed in a normal animal room for 1 month before being tested in a tonal frequency difference detection task. The difficulty of the task was carefully chosen to allow measurement of both improvement and impairment of the performance. The results indicate that discrimination performance was impaired at the exposure frequency and improved at the flanking frequencies (Fig. 3 and Han *et al.*, 2007). Subsequent electrophysiological examination of the primary auditory cortex confirmed that the exposure frequency was over-represented (Fig. 3). Although the impaired discrimination performance for over-represented stimuli is somewhat counter-intuitive, it is consis-

tent with findings that discrimination between prototypical exemplars of the same speech sounds is more difficult than discrimination between non-prototypical exemplars (Kuhl *et al.*, 1992, 2006; Iverson *et al.*, 2003). This phenomenon, also known as the 'perceptual magnet effect', depends on speech sound experience and is specific to the native language (Kuhl, 1991). Thus, it appears that early experience of speech sounds reduces the perceptual sensitivity to subtle differences between prototypical exemplars of the same speech sounds (Kuhl, 1991; Iverson *et al.*, 2003). A computational analysis indicates that the difference detection performance in the sound-exposed animals can be quantitatively accounted for by their cortical frequency representation (Fig. 3 and Han *et al.*, 2007).

Recent probabilistic models of sensory perception suggest that the percept of a stimulus may be shifted towards the stimuli with larger representations (Simoncelli, 2009; Fischer, 2010; Kover & Bao, 2010). In one of the models, all neurons vote for their preferred stimuli and their votes are weighed by their firing rates (Fischer, 2010). The model percept is shifted towards the stimulus with larger representation because more neurons vote for it (Fischer, 2010). Empirical observations support the probabilistic models (Fischer & Pena, 2011; Girshick *et al.*, 2011; Ganguli & Simoncelli, 2014). For example, owls accurately localize sound sources near the center of gaze but systematically underestimate peripheral source directions (Fischer & Pena, 2011). This behavior is correlated with over-representation of the space near the center of gaze and under-representation of lateral space (Fischer & Pena, 2011).

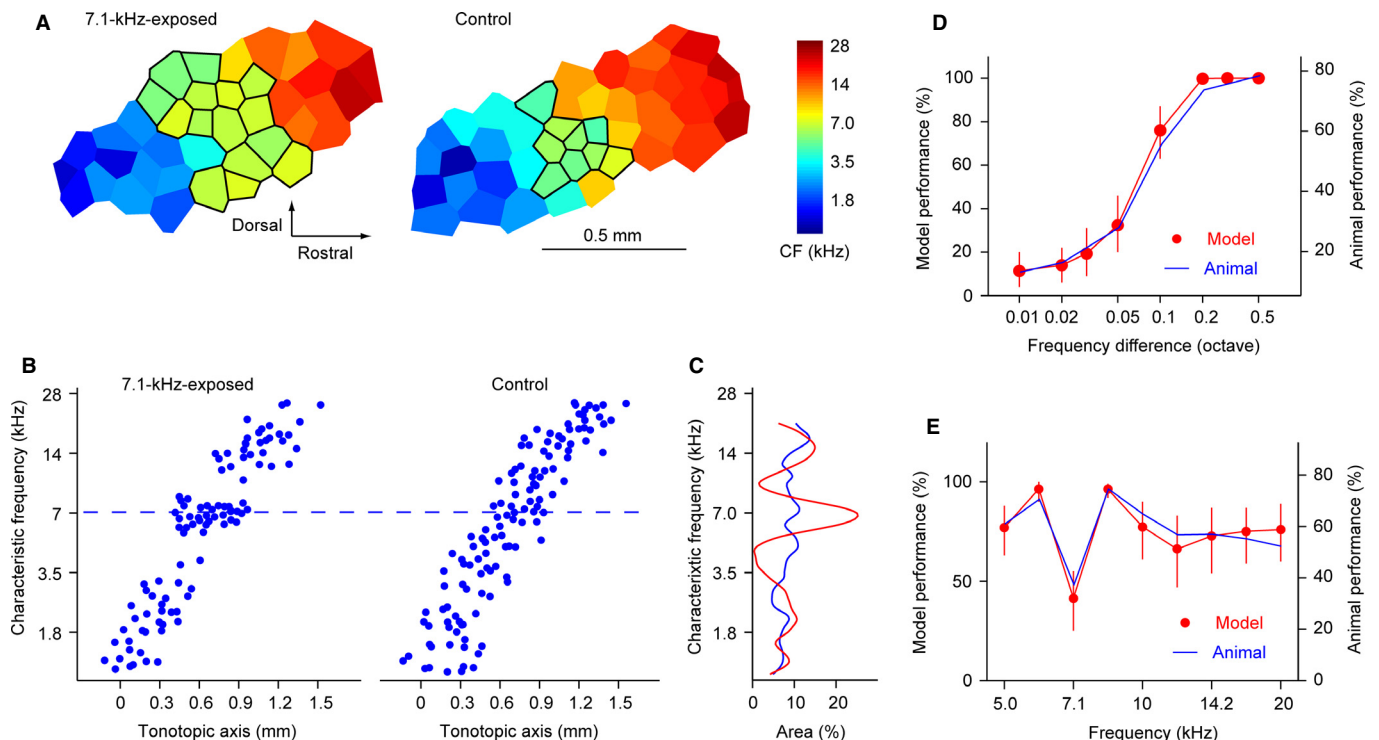


FIG. 3. Experience-dependent cortical plasticity accounts for altered perceptual behaviors. (A) Representative cortical characteristic frequency maps from a 7.1-kHz-exposed animal and a control animal. Neurons in outlined areas had characteristic frequencies in a range of $7.1 \text{ kHz} \pm 0.2$ octave. (B) Characteristic frequency distribution along the tonotopic axis in control and 7.1-kHz-exposed groups. Note the clustering of CFs near 7.1 kHz in the 7.1-kHz-exposed animals. (C) Percentage AI area representing frequencies in a 0.4-octave frequency band. The representations of $7.1 \text{ kHz} \pm 0.2$ octave were significantly larger in tone-exposed animals (red) than in control animals (blue). (D) Comparison of the psychometric function of the model naive AI and that of the naive animals. Note that the performances were scaled for comparison. (E) Comparison of performance of the model 7.1-kHz-exposed AI and that of the 7.1-kHz-exposed animals in the frequency discrimination task. Discrimination of the over-represented frequencies was impaired, and that of under-represented frequencies was improved to nearly the asymptotic 100% level. Error bars in D–E depict 95% confidence interval. This figure was originally published in Han *et al.* (2007).

Higher-order stimulus probability can also shape perception and perceptual behavior. For example, the transitional probability boundary shown in Fig. 2 has been shown to result in a perceptual boundary where difference detection performance is improved (Kover *et al.*, 2013). The altered behavior is correlated with segregated representation of the two frequency bands divided by the probability boundary, and steepened tuning curve slopes at the probability boundary (Kover *et al.*, 2013).

Early experience of natural sounds shapes categorical sensory representation

Above studies suggest that early experience shapes categorical representation and perception of sounds. However, they used simple tone pips. Natural sounds are complex and highly structured. To investigate whether natural sounds shape categorical auditory representation and perception, in a recently published study a group of juvenile rats were exposed to a set of natural animal vocal sounds, referred to as 'jungle sounds' (Bao *et al.*, 2013). The jungle sounds CD loop was an hour of spectrotemporally complex sounds, in which there were at least 40 distinctive repeating motifs of bird songs, mammalian vocalizations and insect sounds (for examples, see Fig. 4A). Cortical neurons became more selective to spectrotemporal

features in the experienced sounds (Bao *et al.*, 2013). At the neuronal population level, more neurons were involved in representing the whole set of complex sounds but fewer neurons actually responded to each individual sound, and with greater firing rates (Bao *et al.*, 2013). A comparison of population-temporal responses to the experienced complex sounds revealed that cortical responses to different renderings of the same song motif were more similar, indicating that the cortical neurons became less sensitive to natural acoustic variations associated with stimulus context and sound renderings (Fig. 4B). By contrast, cortical responses to sounds of different motifs became more distinctive, suggesting that cortical neurons were tuned to the defining features of the experienced sounds. These effects lead to emergent categorical representations of the experienced sounds (Fig. 4C). Further behavioral studies are needed to determine whether the jungle sound-exposure results in behavioral categorical perception of the experienced sounds.

Cortical plasticity is a mechanism for warped and categorical sound perception

Sensory representation and perception have two modes, continuous and categorical, and they serve different purposes (Harnad, 1987). For example, precise localization of sound in space would require continuous and faithful representation of the auditory space. By contrast, recognition of vocalizations or speech sounds would be facilitated by categorical representation. Categorical perception is a result of warped perceptual space, as the perceived difference between stimuli is no longer proportional to their physical difference; stimuli in a categorical center are perceived as being more similar, and stimuli across categorical boundaries are perceived as being more different, than they are (Harnad, 1987). The distortion may allow the system to tune out of irrelevant stimulus variability and noises that otherwise would interfere with the perception. Although categorical readout neurons, those that respond selectively or exclusively to a category of stimuli, are often found in prefrontal cortex (Russ *et al.*, 2007), the underlying neural computation and sensory transformation may happen in the sensory cortex (Ohl *et al.*, 2001; Steinschneider *et al.*, 2003; Chang *et al.*, 2010; Tsunada *et al.*, 2011; Bathellier *et al.*, 2012).

Early sensory exposure results in selective cortical representations of experienced behaviorally important sounds (Zhang *et al.*, 2001; Chang & Merzenich, 2003; de Villiers-Sidani *et al.*, 2008; Zhou *et al.*, 2008; Insanally *et al.*, 2009), and representational boundaries that separate different categories of sounds (Kover *et al.*, 2013). These cortical reorganizations may lead to reduced perceptual sensitivity near the over-represented sounds (Han *et al.*, 2007), shift of percept towards over-represented sounds (Fischer & Pena, 2011) and enhanced perceptual sensitivity at the boundaries of sound categories (Kover *et al.*, 2013). These perceptual distortions are hallmarks of categorical perception. Thus, sensory exposure-induced auditory cortical plasticity during early development may be a mechanism for learning perceptual categories.

Unsupervised learning solves the chicken-and-egg problem in the sensory development

Mammalian sensory systems are remarkably adaptive to the ever-changing environment. This is achieved through different types of learning mechanisms. Supervised and reinforcement learning allow for rapid improvement in sensory processing for stimuli that are behaviorally relevant for the current and specific context. However, these types of learning often require a certain level of sensory

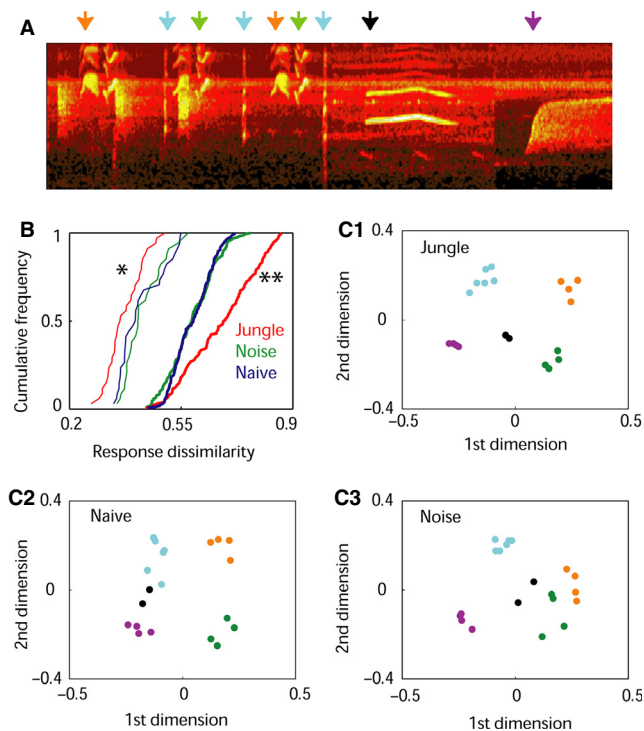


FIG. 4. Segregated representations of jungle song motifs instructed by early experience. (A) Spectrogram of a section of the testing jungle sound stimuli (duration, 4.5 s). Arrows with different colors indicate different song motifs. (B) Cumulative distributions of response dissimilarities. Dissimilarities between responses to different song motifs were plotted with thick lines, and those to same song motifs were plotted with thin lines. (* $P < 0.05$, ** $P < 0.005$ compared with the other two groups). (C) Similarity relations between responses to jungle sounds. Each dot represents a cortical population-temporal response to a jungle sound. The distances between dots are proportional to the dissimilarity between corresponding responses. The color-coding for song motifs is the same as shown in A. Note that cortical responses to different song motifs were sharply segregated in jungle sounds-experienced animals ($P < 0.05$). This figure was originally published in Bao *et al.* (2013).

processing in order to understand instructions or feedback. In addition, the development of the supervised and reinforcement learning mechanisms may also be experience-dependent, and require some level of sensory processing. For example, it is impossible to teach newborns speech in the more traditional and associative form because they cannot understand verbal instructions, and their brain learning systems are still developing.

The sensory system solves this 'chicken-and-egg' problem by unsupervised, statistical learning. Neural circuits in the auditory system organize themselves according to the statistical structures of the sensory input, so that the sounds that are likely to be behaviorally relevant, such as animal vocalizations, are preferentially and categorically represented.

Cortical plasticity underlies perceptual learning

A central question in the field of auditory research is the role of cortical plasticity in perceptual learning. Some studies have found correlations between altered cortical stimulus representations and perceptual learning (Recanzone *et al.*, 1993; Ohl & Scheich, 1996; Bao *et al.*, 2004; Polley *et al.*, 2004, 2006; Schnupp *et al.*, 2006; Han *et al.*, 2007; Froemke *et al.*, 2013), but others have failed to observe those correlations (Talwar & Gerstein, 2001; Brown *et al.*, 2004; Reed *et al.*, 2011; Ranasinghe *et al.*, 2012). As evidence supporting both views accumulates, it is increasingly evident that the conclusion depends on the specifics of the tasks used to measure behavioral performances as well as the types of plasticity effects that are considered (Berlau & Weinberger, 2008). Some auditory tasks preferentially measure procedure or motor learning, and involve corresponding brain substrates of learning. For example, classical conditioning of simple movement to auditory stimuli appears to be mediated by the brainstem and cerebellum (Thompson *et al.*, 1997; Bao *et al.*, 2002), and can be retained even without the forebrain (Mauk & Thompson, 1987). To avoid the confound of non-perceptual influences, tasks that require more sensory processing and less procedure or motor learning are preferred for measuring perceptual learning. Some perceptual tasks measure discrimination of subtle differences between stimuli, while others examine how animals classify very different stimuli. Theoretical and experimental research indicates that fine stimulus discrimination is better achieved by steeper tuning curve slopes, whereas stimulus classification along a large stimulus continuum may be better served by strong responses at the peaks of tuning curves (Butts & Goldman, 2006; Kim & Bao, 2008; Simoncelli, 2009; Kover & Bao, 2010). The multi-dimensional cortical plasticity (e.g., in characteristic frequency, tuning bandwidth, tuning curve slopes, response magnitude, response timing etc.) could have dramatically different effects on different behavioral tasks. Appropriate behavior tests and quantitative models integrating different aspects of cortical plasticity will help elucidate whether and how cortical plasticity influences perception.

Summary

Based on the reviewed evidence, the following hypothetical model emerges. During the critical period of auditory cortical development, sounds with certain properties of natural vocalizations, such as temporal repetition at an ethological rate, profoundly shape acoustic representation in the auditory cortex. Vocalizations that are repeated at the ethological rate in a bout, such as rat pup calls, are represented similarly by a population of neurons. By contrast, different classes of vocalizations that are not produced at the ethological rates in a bout, such as pup and adult rat calls, are represented by separate

populations of neurons. The segregated representations of vocalizations lead to perceptual shifts towards prototypical vocalizations and elevated perceptual contrasts between different classes of vocalizations, resulting in categorical perception of the experienced vocalizations. This model (i) considers categorical perception as a functional outcome of experience-dependent sensory plasticity, (ii) emphasizes the history of sensory experience in understanding how the auditory system represents vocalizations and other ethologically relevant natural sounds and (iii) hypothesizes a causal relationship between sensory representations and perception. Further research that integrates neurophysiological and neuroethological approaches under a quantitative theoretical framework of categorical perception is required to evaluate this model of developmental perceptual learning.

Conflict of interest

The author declares no conflict of interest.

Acknowledgements

The work was supported by National Institute of Deafness and Other Communication Disorders (DC-009259).

Abbreviations

AI, Primary auditory cortex; CF, characteristic frequency.

References

- Bao, S., Chen, L., Kim, J.J. & Thompson, R.F. (2002) Cerebellar cortical inhibition and classical eyeblink conditioning. *Proc. Natl. Acad. Sci. USA*, **99**, 1592–1597.
- Bao, S., Chang, E.F., Woods, J. & Merzenich, M.M. (2004) Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. *Nat. Neurosci.*, **7**, 974–981.
- Bao, S., Chang, E.F., Teng, C.L., Heiser, M.A. & Merzenich, M.M. (2013) Emergent categorical representation of natural, complex sounds resulting from the early post-natal sound environment. *Neuroscience*, **248C**, 30–42.
- Bathellier, B., Ushakova, L. & Rumpel, S. (2012) Discrete neocortical dynamics predict behavioral categorization of sounds. *Neuron*, **76**, 435–449.
- Berlau, K.M. & Weinberger, N.M. (2008) Learning strategy determines auditory cortical plasticity. *Neurobiol. Learn. Mem.*, **89**, 153–166.
- Brown, M., Irvine, D.R. & Park, V.N. (2004) Perceptual learning on an auditory frequency discrimination task by cats: association with changes in primary auditory cortex. *Cereb. Cortex*, **14**, 952–965.
- Butts, D.A. & Goldman, M.S. (2006) Tuning curves, neuronal variability, and sensory coding. *PLoS Biol.*, **4**, e92.
- Chang, E.F. & Merzenich, M.M. (2003) Environmental noise retards auditory cortical development. *Science*, **300**, 498–502.
- Chang, E.F., Rieger, J.W., Johnson, K., Berger, M.S., Barbaro, N.M. & Knight, R.T. (2010) Categorical speech representation in human superior temporal gyrus. *Nat. Neurosci.*, **13**, 1428–1432.
- Domjan, M. 2010. *Principles of Learning and Behavior*. Wadsworth/Cengage, Belmont, CA.
- Ehret, G. & Haack, B. (1981) Categorical perception of mouse pup ultrasound by lactating females. *Die Naturwissenschaften*, **68**, 208–209.
- Elliott, T.M. & Theunissen, F.E. (2009) The modulation transfer function for speech intelligibility. *PLoS Comput. Biol.*, **5**, e1000302.
- Fahle, M. & Poggio, T. (2002). *Perceptual Learning*. MIT Press, Cambridge, MA.
- Fischer, B.J. (2010) Bayesian Estimates from Heterogeneous Population Codes *Proc. IEEE Int'l. Joint Conf. on Neural Networks*.
- Fischer, B.J. & Pena, J.L. (2011) Owl's behavior and neural representation predicted by Bayesian inference. *Nat. Neurosci.*, **14**, 1061–1066.
- Froemke, R.C., Carcea, I., Barker, A.J., Yuan, K., Seybold, B.A., Martins, A.R., Zaika, N., Bernstein, H., Wachs, M., Levis, P.A., Polley, D.B., Merzenich, M.M. & Schreiner, C.E. (2013) Long-term modification of cortical synapses improves sensory perception. *Nat. Neurosci.*, **16**, 79–88.

- Ganguli, D. & Simoncelli, E.P. (2014) Efficient Sensory Encoding and Bayesian Inference with Heterogeneous Neural Populations. *Neural Comput.*, **26**, 2103–2134.
- Garcia-Lazaro, J.A., Ahmed, B. & Schnupp, J.W. (2006) Tuning to natural stimulus dynamics in primary auditory cortex. *Curr. Biol.*, **16**, 264–271.
- Girshick, A.R., Landy, M.S. & Simoncelli, E.P. (2011) Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat. Neurosci.*, **14**, 926–932.
- Han, Y.K., Kover, H., Insanally, M.N., Semerdjian, J.H. & Bao, S. (2007) Early experience impairs perceptual discrimination. *Nat. Neurosci.*, **10**, 1191–1197.
- Harnad, S.R. (1987). *Categorical Perception: The Groundwork of Cognition*. Cambridge University Press, Cambridge.
- Holy, T.E. & Guo, Z. (2005) Ultrasonic songs of male mice. *PLoS Biol.*, **3**, e386.
- Insanally, M.N., Kover, H., Kim, H. & Bao, S. (2009) Feature-dependent sensitive periods in the development of complex sound representation. *J. Neurosci.*, **29**, 5456–5462.
- Iverson, P., Kuhl, P.K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. & Siebert, C. (2003) A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, **87**, B47–B57.
- Keuroghlian, A.S. & Knudsen, E.I. (2007) Adaptive auditory plasticity in developing and adult animals. *Prog. Neurobiol.*, **82**, 109–121.
- Kim, H. & Bao, S. (2008) Distributed representation of perceptual categories in the auditory cortex. *J. Comput. Neurosci.*, **24**, 277–290.
- Kim, H. & Bao, S. (2009) Selective increase in representations of sounds repeated at an ethological rate. *J. Neurosci.*, **29**, 5163–5169.
- Kim, H. & Bao, S. (2013) Experience-dependent overrepresentation of ultrasonic vocalization frequencies in the rat primary auditory cortex. *J. Neurophysiol.*, **110**, 1087–1096.
- Kisilevsky, B.S. & Hains, S.M. (2011) Onset and maturation of fetal heart rate response to the mother's voice over late gestation. *Developmental Sci.*, **14**, 214–223.
- Kisilevsky, B.S., Hains, S.M., Brown, C.A., Lee, C.T., Cowperthwaite, B., Stutzman, S.S., Swansburg, M.L., Lee, K., Xie, X., Huang, H., Ye, H.H., Zhang, K. & Wang, Z. (2009) Fetal sensitivity to properties of maternal speech and language. *Infant Behav. Dev.*, **32**, 59–71.
- Kover, H. & Bao, S. (2010) Cortical plasticity as a mechanism for storing Bayesian priors in sensory perception. *PLoS One*, **5**, e10497.
- Kover, H., Gill, K., Tseng, Y.T. & Bao, S. (2013) Perceptual and neuronal boundary learned from higher-order stimulus probabilities. *J. Neurosci.*, **33**, 3699–3705.
- Kuhl, P.K. (1991) Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Percept. Psychophys.*, **50**, 93–107.
- Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N. & Lindblom, B. (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, **255**, 606–608.
- Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S. & Iverson, P. (2006) Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Sci.*, **9**, F13–F21.
- Liu, R.C., Miller, K.D., Merzenich, M.M. & Schreiner, C.E. (2003) Acoustic variability and distinguishability among mouse ultrasound vocalizations. *J. Acoust. Soc. Am.*, **114**, 3412–3422.
- Mauk, M.D. & Thompson, R.F. (1987) Retention of classically conditioned eyelid responses following acute decerebration. *Brain Res.*, **403**, 89–95.
- Maye, J., Werker, J.F. & Gerken, L. (2002) Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, **82**, B101–B111.
- Nazzi, T., Bertoncini, J. & Mehler, J. (1998) Language discrimination by newborns: toward an understanding of the role of rhythm. *J. Exp. Psychol. Human.*, **24**, 756–766.
- Ohl, F.W. & Scheich, H. (1996) Differential frequency conditioning enhances spectral contrast sensitivity of units in auditory cortex (field AI) of the alert Mongolian gerbil. *Eur. J. Neurosci.*, **8**, 1001–1017.
- Ohl, F.W., Scheich, H. & Freeman, W.J. (2001) Change in pattern of ongoing cortical activity with auditory category learning. *Nature*, **412**, 733–736.
- Partanen, E., Kujala, T., Naatanen, R., Liitola, A., Sambeth, A. & Huotilainen, M. (2013) Learning-induced neural plasticity of speech processing before birth. *Proc. Natl. Acad. Sci. USA*, **110**, 15145–15150.
- Polley, D.B., Heiser, M.A., Blake, D.T., Schreiner, C.E. & Merzenich, M.M. (2004) Associative learning shapes the neural code for stimulus magnitude in primary auditory cortex. *Proc. Natl. Acad. Sci. USA*, **101**, 16351–16356.
- Polley, D.B., Steinberg, E.E. & Merzenich, M.M. (2006) Perceptual learning directs auditory cortical map reorganization through top-down influences. *J. Neurosci.*, **26**, 4970–4982.
- Pons, F. (2006) The effects of distributional learning on rats' sensitivity to phonetic information. *J. Exp. Psychol. Anim. B.*, **32**, 97–101.
- Ranasinghe, K.G., Carraway, R.S., Borland, M.S., Moreno, N.A., Hanacik, E.A., Miller, R.S. & Kilgard, M.P. (2012) Speech discrimination after early exposure to pulsed-noise or speech. *Hearing Res.*, **289**, 1–12.
- Recanzone, G.H., Schreiner, C.E. & Merzenich, M.M. (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.*, **13**, 87–103.
- Reed, A., Riley, J., Carraway, R., Carrasco, A., Perez, C., Jakkamsetti, V. & Kilgard, M.P. (2011) Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron*, **70**, 121–131.
- Rosen, S. (1992) Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos. T. Roy. Soc. B.*, **336**, 367–373.
- Russ, B.E., Lee, Y.S. & Cohen, Y.E. (2007) Neural and behavioral correlates of auditory categorization. *Hearing Res.*, **229**, 204–212.
- Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996) Statistical learning by 8-month-old infants. *Science*, **274**, 1926–1928.
- Saffran, J.R., Johnson, E.K., Aslin, R.N. & Newport, E.L. (1999) Statistical learning of tone sequences by human infants and adults. *Cognition*, **70**, 27–52.
- Sanes, D.H. & Bao, S. (2009) Tuning up the developing auditory CNS. *Curr. Opin. Neurobiol.*, **19**, 188–199.
- Schnupp, J.W., Hall, T.M., Kokelaar, R.F. & Ahmed, B. (2006) Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J. Neurosci.*, **26**, 4785–4795.
- Schreiner, C.E. & Polley, D.B. (2014) Auditory map plasticity: diversity in causes and consequences. *Curr. Opin. Neurobiol.*, **24**, 143–156.
- Simoncelli, E.P. (2009) Optimal estimation in sensory systems. In Gazzaniga, M. (Ed.), *The Cognitive Neurosciences*. MIT Press, Cambridge, MA, pp. 525–535.
- Singh, N.C. & Theunissen, F.E. (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. *J. Acoust. Soc. Am.*, **114**, 3394–3411.
- Steinschneider, M., Fishman, Y.I. & Arezzo, J.C. (2003) Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *J. Acoust. Soc. Am.*, **114**, 307–321.
- Talwar, S.K. & Gerstein, G.L. (2001) Reorganization in awake rat auditory cortex by local microstimulation and its effect on frequency-discrimination behavior. *J. Neurophysiol.*, **86**, 1555–1572.
- Thompson, R.F., Bao, S., Chen, L., Cipriano, B.D., Grethe, J.S., Kim, J.J., Thompson, J.K., Tracy, J.A., Weninger, M.S. & Krupa, D.J. (1997) Associative learning. *Int. Rev. Neurobiol.*, **41**, 151–189.
- Toro, J.M. & Trobalon, J.B. (2005) Statistical computations over a speech stream in a rodent. *Percept Psychophys.*, **67**, 867–875.
- Tsunada, J., Lee, J.H. & Cohen, Y.E. (2011) Representation of speech categories in the primate auditory cortex. *J. Neurophysiol.*, **105**, 2634–2646.
- de Villers-Sidani, E., Simpson, K.L., Lu, Y.F., Lin, R.C. & Merzenich, M.M. (2008) Manipulating critical period closure across different sectors of the primary auditory cortex. *Nat. Neurosci.*, **11**, 957–965.
- Vouloumanos, A. & Werker, J.F. (2007) Listening to language at birth: evidence for a bias for speech in neonates. *Developmental Sci.*, **10**, 159–164.
- Wang, X. & Kadia, S.C. (2001) Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J. Neurophysiol.*, **86**, 2616–2620.
- Zhang, L.L., Bao, S. & Merzenich, M.M. (2001) Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nat. Neurosci.*, **4**, 1123–1130.
- Zhou, X., Nagarajan, N., Mossop, B.J. & Merzenich, M.M. (2008) Influences of un-modulated acoustic inputs on functional maturation and critical-period plasticity of the primary auditory cortex. *Neuroscience*, **154**, 390–396.