

Opinion

Unraveling the Biology of Auditory Learning: A Cognitive–Sensorimotor–Reward Framework

Nina Kraus^{1,2,3,4,*} and Travis White-Schwoch^{1,4}

The auditory system is stunning in its capacity for change: a single neuron can modulate its tuning in minutes. Here we articulate a conceptual framework to understand the biology of auditory learning where an animal must engage cognitive, sensorimotor, and reward systems to spark neural remodeling. Central to our framework is a consideration of the auditory system as an integrated whole that interacts with other circuits to guide and refine life in sound. Despite our emphasis on the auditory system, these principles may apply across the nervous system. Understanding neuroplastic changes in both normal and impaired sensory systems guides strategies to improve everyday communication.

Learning, Language, and Communication

Nervous system plasticity has been observed across the animal kingdom from single cells to sophisticated circuits. Sensory systems are prodigious in their ability to reshape response properties following learning, and in the auditory system plasticity has been observed from cochlea to cortex. This learning is fundamental to our ability to function in and adapt to our environments. Experience navigating this sensory world drives language development – perhaps the most remarkable auditory learning task humans accomplish – and it is necessary to understand the principles that govern this plasticity to devise strategies to improve language and communication in normal and disordered systems.

Here we argue that cognitive, sensorimotor, and reward ingredients engender biological changes in sound processing. The mechanisms behind these changes lie in two sets of dichotomous systems: (i) the afferent projections that transmit signals from ear to brain and the efferent projections that propagate signals from brain to ear; and (ii) the primary and nonprimary processing streams that suffuse the auditory neuraxis (Figure 1). We highlight experiments that advance our understanding of the neurophysiological foundations underlying **auditory processing** (see Glossary) and that offer objective markers of auditory processing in humans. Finally, we place learning in the context of a distributed, but integrated, auditory system.

Rethinking the Auditory System: A Distributed, but Integrated, Circuit

Traditional models characterized the auditory system as series of relay stations along an assembly line, each with distinct functions [1–3]. While these hierarchical models recognized the interconnectivity of the system, the emphasis was to characterize each nucleus's specialization. The idea was that understanding each station would build each block necessary to construct the auditory circuit, and this 'inside-out' approach has contributed greatly to our understanding of auditory neurophysiology.

Trends

The auditory system should be thought of as a distributed, but integrated, circuit that is more than a simple set of processing stations.

Experiences sculpt the auditory system and impart a biological 'memory' that can change automatic response properties from cochlea to cortex.

The cognitive, sensorimotor, and reward aspects of these experiences optimize auditory learning.

¹Auditory Neuroscience Laboratory and Department of Communication Sciences, Northwestern University, Evanston, IL, USA

²Department of Neurobiology and Physiology, Northwestern University, Evanston, IL, USA

³Department of Otolaryngology, Northwestern University, Chicago, IL, USA

⁴These authors contributed equally to this work.

*Correspondence: nkraus@northwestern.edu (N. Kraus).

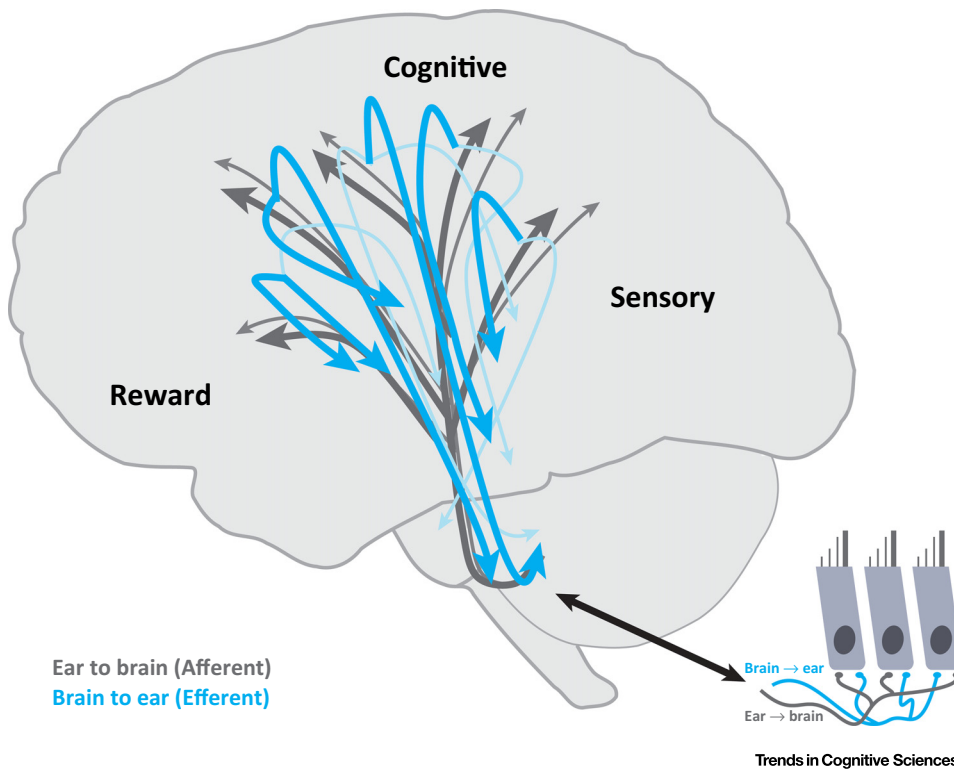


Figure 1. The Auditory System is a Distributed, but Integrated, Circuit. Key to this framework is the rich series of afferent (ear-to-brain/bottom-up) and efferent (brain-to-ear/top-down) projections that pervade every station along the auditory pathway, including to and from the cochlear hair cells (inset). These pathways contain primary (darker colors) and nonprimary (lighter colors) divisions of the auditory system and facilitate both sound processing and neural plasticity. Successful auditory learning engages cognitive, sensorimotor, and reward networks and the intersection of these circuits guides neuroplasticity.

We propose a complementary ‘outside-in’ approach. Our view is that the auditory system should be thought of as a distributed, but integrated, circuit (Figure 1). Any acoustic event catalyzes activity throughout the **auditory neuraxis** and we argue that sound processing – and any assay thereof – is a reflection of this integrated network. Although each structure is specialized to perform a specific function, this specialization has evolved in the context of the entire circuit. To understand auditory learning, then, we are forced to move past a focus on an individual processing station as a single locus of activity, expertise, or disorder.

Our view is consistent with an emerging trend in neuroscience to consider the interplay of multiple processing stations and the ‘give and take’ between cortical and/or subcortical systems underlying human behavior [4–8].

Plasticity in the Human Auditory System: A Double-Edged Sword

We regard everyday auditory experience as a learning process that shapes the nervous system, not least because auditory experience is necessary for the maturation of basic auditory circuits [9–11,106]. These changes may be exacerbated – for better or worse – and cases of expertise and deprivation both contribute to understanding how experience shapes auditory circuitry [12]. Neuroplasticity must therefore be viewed as a double-edged sword. The cognitive, sensorimotor, and reward ingredients of auditory experience drive plasticity and a hypothesis based on this framework is that insults to any of these domains dictate the resulting phenotype.

Glossary

Auditory neuraxis: the auditory information processing pathway of the nervous system that transmits information back and forth between the cochlea and cortex.

Auditory processing: a cluster of listening skills that refers to the ability to make meaning from sound. Listeners can have normal hearing thresholds but still struggle to process auditory information.

Frequency-following response (FFR)/auditory brainstem response to complex sounds (cABR): a scalp-recorded potential that comprises aggregate neural processing of sound details and that captures a snapshot of the integrity of auditory processing.

While historically ‘FFR’ referred to responses to low-frequency pure tones, the FFR can be as rich and complex as the eliciting stimulus, and we use it to refer to neural activity that ‘follows’ both transient and periodic acoustic events.

Inhibitory control: the ability to actively suppress information irrelevant to the task at hand.

Otoacoustic emissions: sounds generated by the outer hair cells of the inner ear; in certain cases these sounds can be modulated by active listening.

Phase locking: the ability of auditory neurons to change their intrinsic rhythms to follow those of incoming sounds.

Phoneme: the smallest unit of speech that conveys a change in meaning. Phonemic information is connoted by fine-grained acoustic contrasts. For example, the acoustic difference between /b/ and /g/ is phonemically meaningful, but the acoustic difference between /p/ in [putter] and /p/ in [sputter] is not.

Statistical learning: an implicit process of picking up on the statistical regularities in the environment; infants exhibit this ability and it is thought to be a principal component of language learning.

Box 1. Indexing Auditory Processing in Humans: A Matter of Time

The acoustic world unfolds at once across timescales, from subsecond syllables to multiminute monologues. Much as the visual system must integrate basic cues such as color, edge, and motion into a coherent object, the auditory system must integrate acoustic cues across time and frequency into meaningful percepts. It has been argued that insensitivity to temporal cues at one or more rates may contribute to language impairment [91,94] and so an important goal is to understand how the brain makes sense of information within and across these timescales.

Neurophysiological responses to speech sounds provide glimpses of the integrity of the processing of key acoustic features across timescales. The FFR reflects neural processing of fast acoustic information such as subsyllabic cues. A major advantage of this approach is the physical symmetry between the evoking stimulus and the response (Figure 1), meaning that the latter reflects the integrity with which any acoustic cue is transcribed: consonants and vowels, prosody, timing, pitch and harmonics, and more. Thus, within a single evoked response rests a plethora of information about how well details of sound are coded. Indeed, when the FFR is played through a speaker it is recognizable as the eliciting stimulus [95].

Moreover, FFR properties are linked to everyday listening skills. Few of these are as complex and computationally demanding as understanding speech in noise, which depends on a series of interactions between factors both exogenous (the talker, his or her accent, the language being spoken, and the acoustics of the room and noise) and endogenous (the listener and his or her experience, cognitive abilities, and hearing acuity). Due to these demands – particularly the demands for speed and precision in auditory processing – it stands to reason that any number of insults may constrain these processes; indeed many clinical populations exhibit difficulties recognizing speech in noise. In this regard, the ability to recognize speech in noise may reflect overall brain health. FFR properties are linked to these listening challenges, suggesting that it may be an approach to uncover individual differences in listening abilities and their responses to intervention [96–100], thereby providing a biological indication of brain health. FFR is agnostic to a subject's age and species: the same protocols have been used as early as infancy [101], across the lifespan [102], and in animal models [103], presenting granularity and uniformity to the study of sound processing. Thus, it can provide an approach to inform links between neural function and everyday communication such as hearing speech in noise.

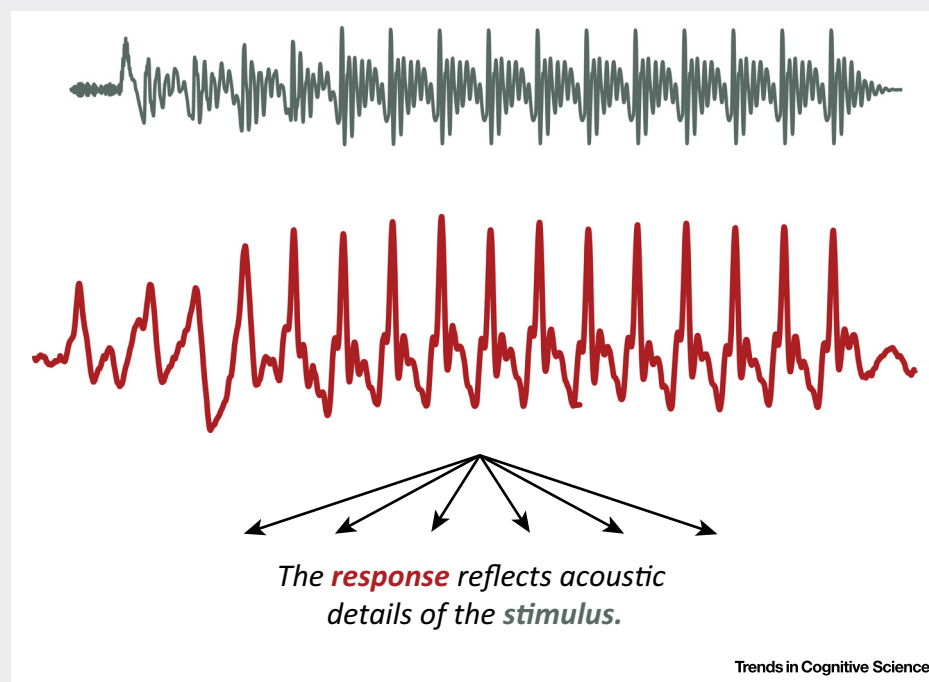
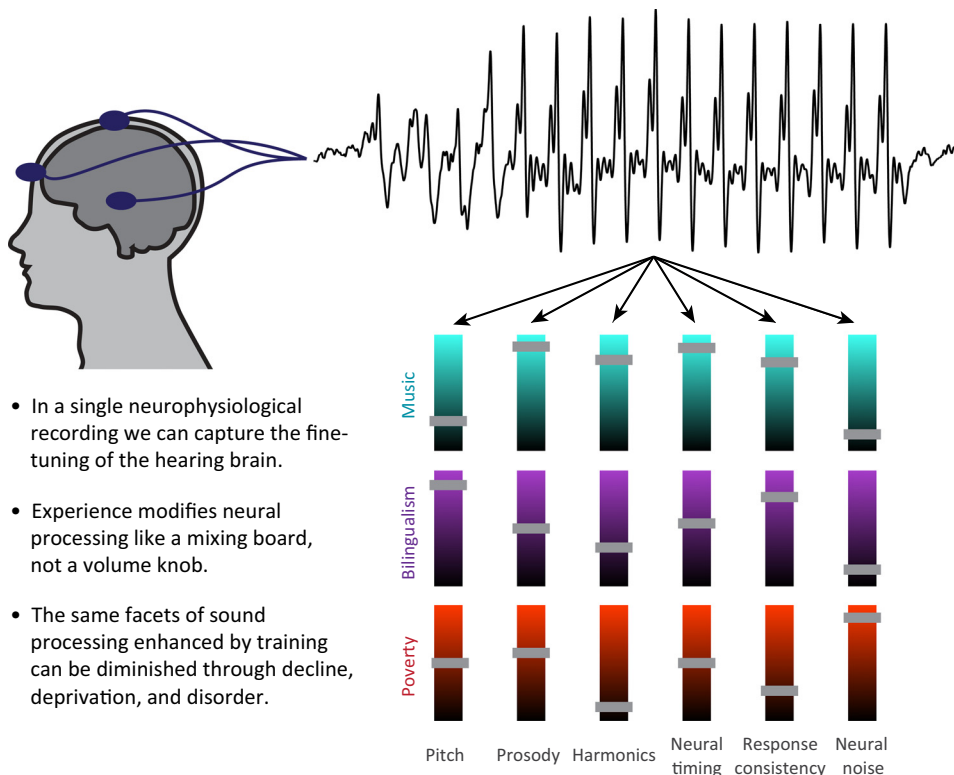


Figure 1. The FFR is as Complex as the Eliciting Stimulus and Mirrors Many of the Physical Properties.

Of particular interest in our research program is the neural coding of fast auditory events, such as the details that convey **phonemic** identity in speech (Box 1). Our laboratory has developed an approach to index the influence of life experience on the neural coding of these fast acoustic details called the **frequency-following response (FFR)**. We have previously referred to this as the **auditory brainstem response to complex sounds (cABR)**, but fear that this terminology

Key Figure

Measuring Neural Responses to Speech Allows Us to Evaluate Auditory Processing – and the Legacy of Auditory Experience – in Humans



Trends in Cognitive Sciences

Figure 2. Scalp electrodes pick up neural firing in response to sound and the brainwave recapitulates a life in sound by reflecting the fine-tuning of the hearing brain through experience. The nature of an individual experience shapes the nature of the plasticity: different elements of sound processing are selectively modulated, for better or worse, within an individual. This is illustrated through a mixing-board analogy, with several aspects of sound processing illustrated; the short bars reflect enhancements (above midline) or diminutions (below midline) of auditory processing. Bars at the halfway point reflect aspects of sound processing that appear unaffected by that particular experience. Although we highlight several aspects of sound processing in this illustration, much more may be glimpsed through these neurophysiological responses.

undermines the integrated and experience-dependent nature of the activity it indexes. The FFR is as complex as the eliciting stimulus and we use ‘FFR’ to refer to the product of aggregate neural activity in the auditory midbrain that reflects the coding of aggregate speech features, including activity that ‘follows’ both transient and static acoustic cues; because the auditory midbrain is a ‘hub’ of intersecting afferent and efferent auditory projections, in addition to projections to and from non-auditory cortices, its response properties are shaped by this constellation of cognitive, sensory, and reward input (Figure 2, Key Figure). Thus, despite its subcortical basis, the FFR reflects the distributed, but integrated, activity emblematic of auditory processing.

This research emphasizes the imprint of changes to the auditory system that affect the automatic sound processing that is always on and cannot volitionally be turned off, even after training has stopped [13,14]. Thus, biological infrastructure in the auditory system is influenced by an

individual's life in sound. No two people hear the world exactly the same way because acoustic experiences impart enduring biological legacies (Figure 2).

Cognitive Influences on Auditory Processing

The cognitive component of our framework is grounded in these principles: (i) listening is an active process that engages cognitive networks; (ii) the precision of automatic sound processing in the brain is linked to cognitive skills such as attention and working memory; and (iii) the cognitive systems engaged during listening selectively modulate the aspects of sound that are behaviorally relevant. The legacy of this repeated, active engagement is engrained in the nervous system over time as listeners make sound-to-meaning connections.

Speech understanding relies on the ability to pull on cognitive functions such as working memory and attention [15–18]; engagement of these systems strengthens the neural circuits that facilitate listening [19]. One study showed that cognitive factors shape auditory learning in an experiment comparing two groups of rats [20]. The first group trained to attend to frequency contrasts in a set of tones, whereas a different group trained on intensity contrasts; crucially, identical stimulus sets were used in the two groups. Cortical maps changed along the acoustic dimension that was trained, demonstrating that what is relevant to an animal dictates map plasticity (see also [107]).

In humans, several studies show links between the integrity of the neural processing of sound and cognitive abilities [21], suggesting that the legacy of cognitive engagement is revealed through the precision of neural function. Additionally, this suggests that training to strengthen a cognitive skill propagates to sensory systems [22–24].

Many of these insights come from studies of music training, which provides a model to understand the biology of auditory learning [25–27]. Making music requires an individual to engage multiple cognitive systems and to direct attention to the sounds that are heard, produced, and manipulated. The physical act of producing sound – through instrument or voice – mandates intricate motor control and stimulates auditory–motor projections [28]. In addition, music is an inherently rewarding stimulus that elicits activity throughout the limbic system [29]. The musician's brain has been finely tuned to process sound and the musician is a case to explore what is possible in terms of experience-dependent plasticity.

Regarding cognitive–sensory coupling, individuals with music training exhibit stronger neural coding of speech in noise concomitant with heightened auditory working memory [30]. Contrast this with a bilingual, who exhibits stronger neural coding of pitch cues concomitant with heightened **inhibitory control** [31]. A musician needs to pull out another instrument's 'voice' from an ensemble while mentally rehearsing a musical excerpt, facilitating the processing of signals in a complex soundscape and exercising working memory. However, a bilingual needs to actively suppress one mental lexicon while using voice pitch as a cue to activate the appropriate one. Whereas music training is associated with superior speech recognition in certain types of background noise [32,33] (cf. [34]), the cognitive systems engaged through bilingualism create a different situation. Bilinguals have superior recognition of non-speech sounds in noise but inferior recognition of speech in noise, due to cognitive interference from the mental lexicon they are attempting to suppress during active listening [35,36]. Thus, the impact of this cognitive–sensory coupling for everyday listening skills depends on what constellation of cognitive and sensory skills are rehearsed. This juxtaposition illustrates an important principle of auditory plasticity: cognitive systems tune into particular details of sound and selectively modulate the sensory systems that represent those features (Figure 2). By analogy, then, auditory learning may be thought of as a 'mixing board' more than a single 'volume knob', with distinct aspects of neural coding selectively modulated as a function of the precipitating experience [37,38]. This contrast

also reinforces the notion of a double-edged sword in experience-dependent plasticity and adds a layer of nuance: within an individual some listening skills may be strengthened whereas others may be suppressed.

There is similarly a tight interplay between cognitive and sensory losses; older adults with hearing loss exhibit faster declines in working memory, presumably because degraded auditory acuity limits opportunities for cognitively engaging and socially rewarding interactions [39]. Training these cognitive skills, however, cascades to boosts in sensory processing. For example, older adults have delayed neural timing in response to consonants but not vowels [40]; auditory–cognitive training that directs attention to consonants (including built-in reward cues) reverses this signature aging effect [23]. Similar phenomena are observed following cognitive interventions in the visual system [24,41].

These studies illustrate that identical neural pathways are imputed in disorder and its remediation and are consistent with the view that both should be conceptualized as auditory learning. They demonstrate how fine-grained aspects of sound processing are selectively modulated based on the cognitive demands and bottlenecks of the experience (Figure 2). Moreover, these cases exemplify the coupling between the integrity with which the nervous system transcribes sound and the cognitive skills important for everyday listening.

Sensorimotor Influences on Auditory Processing

The sensorimotor component of our framework is grounded in these principles: (i) the infrastructure responsible for encoding basic sound features is labile; (ii) extreme cases of deprivation and expertise illuminate mechanisms that apply to a typical system; and (iii) the entire auditory pathway – including the hair cells – can be thought of as sites of ‘memory storage’ because response properties reflect the legacy of auditory experience(s).

Basic sensory infrastructure has a potential for reorganization. The most extreme examples comprise cases of profound deprivation, such as deafness, blindness, or amputation, where sensory cortices are coopted by circuits dedicated to the remaining senses, but only after a period of adaptation (that is, learning) [42,43,108]. These extremes illustrate the brain's potential for reorganization and the mechanisms underlying this remodeling.

In terms of expertise, music again offers a model for auditory learning. Musicians process sound more efficiently even when not playing music, suggesting that repeated active engagement with sound shapes the automatic state of the nervous system [19]. The imprint of music training extends all the way to the outer hair cells of the cochlea [44,45]. The musician model also demonstrates that sensory input alone is insufficient to drive neural remodeling: comparisons between children undergoing active music training (that engages cognitive, motor, and reward networks) and those in music appreciation classes have shown neurophysiological changes only in the former [46]. Thus, sensory input may be necessary, but not sufficient, for auditory learning [47].

With regards to language learning, evidence from songbirds demonstrates a causal role for the basal ganglia in song learning [48], suggesting a role for the motor system in language learning. We are just beginning to learn how the motor system is involved in auditory learning in humans, but it seems that motor acuity is tied to language abilities [49,50] and that training rhythmic skills can boost literacy skills [51]. The rhythm–language link may underlie the observation that music training confers gains in reading achievement.

Finally, we mention an example of sensory learning that on its surface appears to occur automatically. Infants quickly learn statistical regularities in the acoustic environment and this is thought to contribute to language acquisition [52]. However, not even these ostensibly passive

learning processes are exempt from cognitive influence: prior experience and active expectations guide **statistical learning** [53,54]. Thus, as young as infancy, listeners can connect incoming sounds to meaning and also exert meaning on incoming sounds, reinforcing the interplay between sensorimotor and cognitive systems in auditory learning.

Reward (Limbic) Influences on Auditory Processing

The reward component of our framework is grounded in these principles: (i) reward systems spark reorganization in fundamental auditory infrastructure; (ii) social and reward contexts gate auditory learning in humans; but (iii) limbic input can create the conditions to learn something that does not optimize auditory processing.

We learn what we care about. Consequently, the limbic system is likely to facilitate neural remodeling. Classic studies show that stimulation of the cholinergic nucleus basalis galvanizes cortical map reorganization [55,56] see also [109,110]. Aberrant sensory–limbic coupling, in turn, is involved in disorders such as tinnitus [57], but also in their treatment [47]. This again emphasizes that identical networks are implicated in conditions of both enhancement and diminution of sound processing.

Less is known about how the limbic system guides auditory learning in humans, in part due to practical limitations in controlling the expression of neuromodulators (although early evidence is promising [58]). Once again, music training provides a model: listening to and producing music activates multiple auditory–limbic projections [29,59]. Given that music training directs attention to minute details of sound in a rewarding context, it stands to reason that these neuromodulators play a role in the resulting neural remodeling.

The limbic system may also play a role in language development. It has been argued that infants must tune into the aforementioned statistical patterns in the auditory environment to jumpstart language learning, but that these computations are gated by social (i.e., reward) context [60]. For example, infants learn non-native phonemic contrasts when they are modeled by a tutor speaking ‘motherese’, but only if that tutor is present and interacting with the child – a video of the tutor is insufficient [61].

Deficits in reward input, then, are hypothesized to contribute to language impairment. Children with autism, for example, exhibit reduced functional connectivity between limbic structures and voice-selective auditory cortex, which suggests a decoupling of sensorimotor and reward networks during everyday listening [62]. Indeed, many children with autism show poor neural coding of prosodic cues that convey emotional and pragmatic information [63].

Children whose mothers have relatively low levels of education – a proxy for socioeconomic status – present a different case of deprivation. Children in these homes hear approximately 30 million fewer words than their peers; in addition, they hear two-fifths of the number of different words, meaning that both the quantity and the quality of their everyday linguistic input is impoverished [64]. Consider that a mother's voice is perhaps the single most rewarding sensory cue available to a child. If the sensory input is impoverished, but the conditions are right for learning, what is learned may itself be impoverished. This linguistic impoverishment is reflected by poor neural coding and cognitive skills [65] (Figure 2). This is consistent with evidence from animal models that environmental deprivation constrains nervous system development; environmental enrichment, however, reverses this maladaptive plasticity [66], reinforcing the concept of auditory learning as a double-edged sword. This hypothesis also aligns with evidence that task reward structure shapes not only whether plasticity occurs, but how it manifests [67,68].

Taken together, these studies illustrate that, on the one hand, a lack of reward structure stymies the mechanisms of auditory learning. On the other hand, sufficient reward structure with an

impoverished content may cause learning of the wrong material. Presumably, this principle applies to the cognitive and sensorimotor aspects of auditory learning as well.

Mechanisms of Learning

Having laid the groundwork to understand that cognitive, sensorimotor, and reward systems are necessary to drive neural remodeling, the question arises: how do these systems influence automatic sound processing?

Two anatomical dichotomies help navigate this integrated circuit and its role in auditory learning: afferent versus efferent projections and the primary versus nonprimary pathways.

Afferent and Efferent Projections: Bottom-Up Meets Top-Down

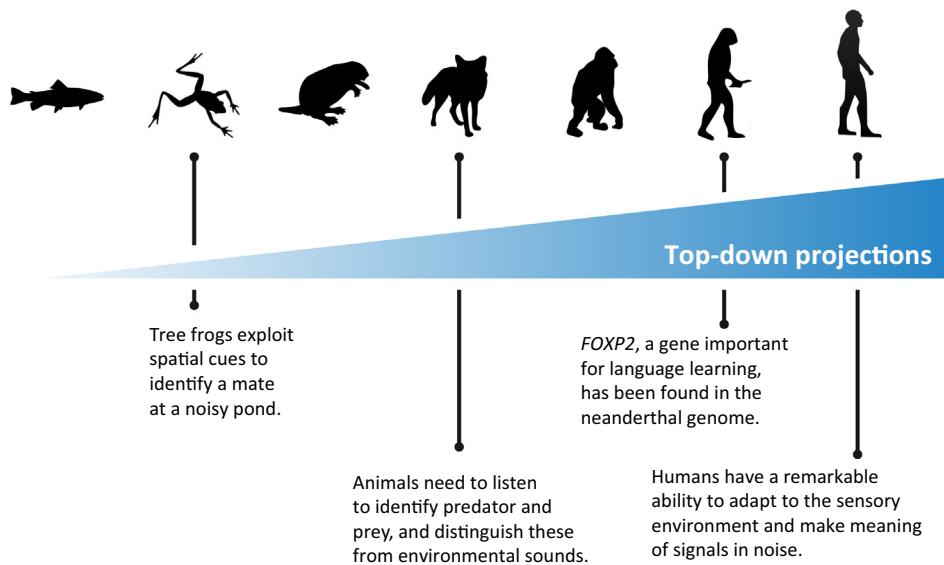
The first dichotomy comprises the projections that feed signals forwards and backwards through the auditory system (Figure 1). The bottom-up afferent projections transmit information forwards to accomplish signal processing ('ear to brain') whereas the top-down efferent projections propagate signals backwards ('brain to ear'); both extend between the cochlea and cortex [69] and the latter mediates remodeling in subcortical structures [70,71].

Our proposal is that the efferent network shapes automatic response properties in cochlear and subcortical systems, which is why the basic response properties of the auditory system, such as **otoacoustic emissions** and electrophysiological responses, reflect life experiences in sound. It has been argued that similar mechanisms underlie both attention-driven online changes and long-term plasticity [67]. This leads to the hypothesis that if these experiences that engage cognitive, sensorimotor, and reward systems are repeated sufficiently they can, over time, facilitate functional remodeling by imparting a 'memory' to afferent processing [19,72] and future learning [73–75].

We hypothesize that the efferent system has become larger and more intricate evolutionarily with increasingly sophisticated auditory behaviors. Numerous complex auditory behaviors – many of which are important for listening in everyday situations – are similar across species. This includes the learning observed in animals with precocious auditory abilities such as bats [76], ferrets [77], and humans [78]. This may be due to convergent evolution, the independent evolution of a trait in distinct lineages based on the needs of the organism. These behaviors are perhaps most sophisticated in humans (Figure 3) and we speculate that the convergent evolution of efferent projections may underlie some of these behaviors and the key role that auditory learning plays in developing the skills necessary for effective everyday communication. If one accepts that language learning pulls on the circuitry necessary for auditory learning, one could imagine a role for the efferent system in language development and poor activation of these top-down networks as a chief factor in language impairment [79].

Primary and Nonprimary Divisions

The second dichotomy pertains to auditory structures such as the cochlear nucleus, inferior colliculus, thalamus, and cortex: the distinction between primary and nonprimary pathways (also known as lemniscal and paralemniscal, cochleotopic and diffuse, or highway and country road; Figure 4). Neurons in the primary pathway are biased to respond to auditory stimuli, whereas the nonprimary neurons are more multisensory. The primary pathway is tonotopically organized, shows sharp tuning, and strongly **phase locks** to the stimulus, whereas the nonprimary pathway is not especially tonotopic, has broader tuning, and does not time lock as strongly to stimuli [80]. Partly for these reasons, it is thought that the primary auditory cortex ('core') represents nearly all incoming signals whereas the nonprimary cortex ('belt' and 'parabelt') specializes for communication signals such as speech and music [81–83].

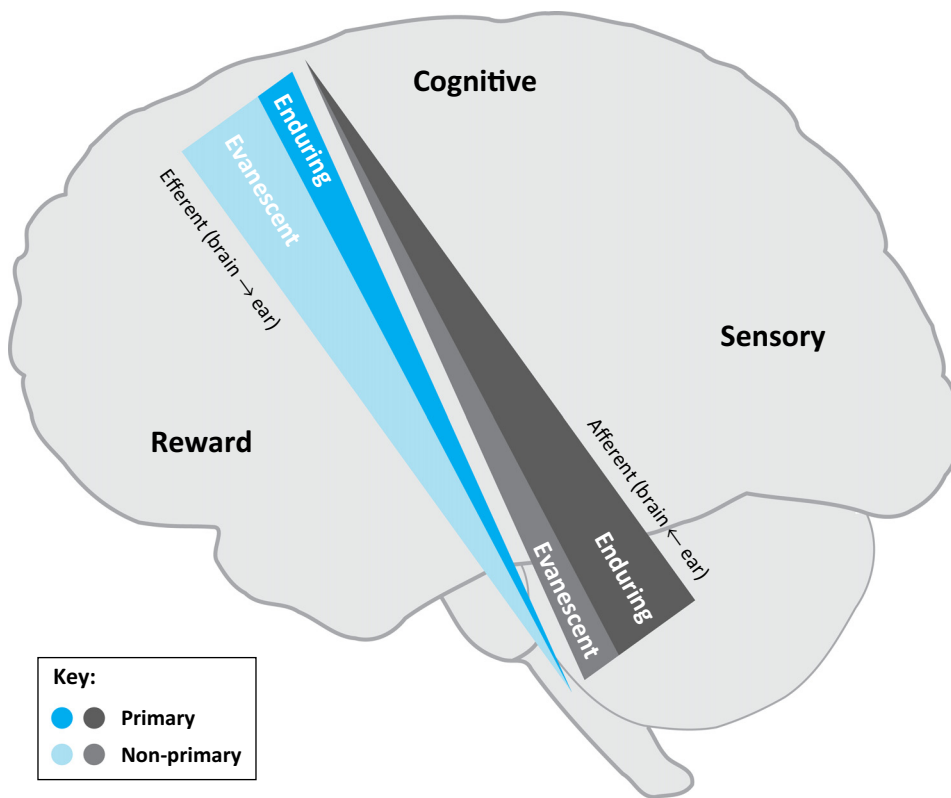


Trends in Cognitive Sciences

Figure 3. We Propose That the Corticofugal System (Top-Down Projections) Has Become Richer Evolutionarily, with a Larger Number of, and More Connections Between, Fibers with Increasing Phylogenetic Sophistication. We speculate that this underlies some of the increasingly sophisticated behaviors observed across species. These behaviors are likely to have emerged convergently; that is, they evolved independently in distinct lineages as a function of the organism's communication needs. Frogs are capable of exploiting many of the basic acoustic cues we use in complex soundscapes, such as spatial hearing and listening in dips in background noise [104]. More sophisticated animals had to make meaning through diverse environmental sounds, learning both to ignore the rustling wind and to hustle when a predator approached. We propose that language learning is contingent on the rigor and activation of this system, and it is interesting to note that our close genetic ancestor *Homo neanderthalensis* carried *FOXP2* [105], a gene implicated in language learning and impairment.

Adopting our systems-wide perspective, however, an additional distinction emerges: the primary processing stream preferentially codes fast temporal information whereas the non-primary stream codes relatively slow information [84–86]. This hypothesis is consistent with evidence from the rat trigeminal system [87] and primate visual system [88] that parallel pathways code fast and slow information. The functional consequences for language development are only beginning to be understood; however, preliminary evidence suggests that deficits in either fast or slow auditory–temporal processing may lead to language impairment, but may not necessarily co-occur [89–91].

Less is known with regard to learning and the primary versus nonprimary pathways. The nonprimary pathway has been implicated in rapid task-related plasticity, such as adapting to stimulus context [92] and classical conditioning [93]. During active listening, neurons in the prefrontal cortex first change their tuning, followed by neurons in the nonprimary auditory cortex, and then finally neurons in the primary auditory cortex [7]. This leads to the hypothesis that the nonprimary system is more labile than the primary circuitry and may facilitate rapid online learning and adaptation in connection to cognitive and reward circuits. We speculate that changes to the afferent pathway are biased towards stability: the system exhibits a more enduring physiology that resists transient changes and relatively few of the projections exhibit task-related evanescence (Figure 4). Conversely, the efferent pathway is biased towards in-the-moment changes; this evanescence facilitates phenomena such as selective attention to one speaker. This system is relatively less persistent in sound processing. The more an activity is done, repeated, and over-learned, the more likely remodeling will occur in the primary pathway and, eventually, influence afferent processing.



Trends in Cognitive Sciences

Figure 4. The Primary and Nonprimary Pathways Operate in Parallel throughout the Auditory System. Each is schematized as a wedge. The larger ear-to-brain wedge (afferent primary) illustrates the predominantly enduring automatic processing whereas the smaller ear-to-brain wedge (afferent nonprimary) illustrates the relatively smaller degree of evanescence. This dichotomy is flipped in the efferent system, where the larger wedge (efferent nonprimary) shows a predominance of evanescence in processing whereas the smaller wedge (efferent primary) suggests that this system is relatively less stable. The trade-off between stability and evanescence between the afferent and efferent systems may underlie the ability to maintain enough plasticity to adapt to new situations while also retaining enough stability to pull on previous experiences (language, memory, knowledge of the sensory world). The more an auditory activity is performed, repeated, and over-learned, it transfers to the primary pathway, which becomes a repository of auditory experience by virtue of changes to its basic response properties.

This primary-enduring, nonprimary-evanescent distinction may underlie the capacity to strike a balance between the stability of auditory processing and malleability in attention, adaptation, and learning (see Outstanding Questions).

Concluding Remarks and Future Directions

We have reviewed the auditory system's ability to change. In particular, we have argued that cognitive, sensorimotor, and reward systems optimize auditory learning and that this learning underlies success in everyday language and communication. We have also argued that the auditory system should be thought of as a distributed, but integrated, circuit that is a moving target; for better or worse, its response properties change through the interplay of cognitive and reward circuits during everyday listening. Thus, both expertise and disorder should be considered from a common standpoint of neuroplasticity. While our emphasis has been on the auditory system, we argue that these principles extend to other sensory systems [4,6,8].

The recognition that states of decline, deprivation, and disorder should be viewed through a lens of plasticity suggests that they may, in part, be reversible. If the same pathways are responsible

Outstanding Questions

How do auditory experiences layer and interact across an individual's life? How does attention in the past facilitate or constrain future learning?

How do the indices of learning discovered in subdivisions of the auditory pathway work together as functional processing of sound becomes shaped by experiences?

Is the neural processing of particular sound details more or less malleable with experience?

How does the auditory system balance temporal processing across timescales of acoustic information? Is there a single 'timekeeper' of nested oscillators or is each distinct? Does plasticity at one timescale of auditory processing imply plasticity at multiple timescales?

What 'dosage' of auditory training is necessary to impart meaningful and lasting neurophysiological and behavioral changes? Our framework would predict that training combining cognitive, sensorimotor, and reward components would be optimized for fast and long-lasting changes.

Can evolutionary, comparative studies of the corticofugal pathway explain the phenomenal learning capacity observed in acoustically sophisticated species? Is this interconnectedness at the heart of language and auditory learning? We speculate that the ability to learn and modulate sensory infrastructure has increased evolutionarily and thus the influence of auditory learning on everyday behavior is greater in more sophisticated species (see Figure 3 in main text).

How does our framework for auditory processing extend to other sensory systems?

How can the lessons of auditory learning be transferred outside the laboratory and into clinical, educational, and community settings?

for expertise and disorder, the conditions that facilitate expertise may ameliorate communicative difficulties. Our framework therefore makes a clear case for auditory training as an intervention for listening and language difficulties, and – providing the training integrates cognitive, sensorimotor, and reward systems – early evidence is promising.

Finally, we have highlighted how measuring the integrity of sound processing at basic levels of the auditory system opens a window on human communication and the imprint of a life spent in sound. A healthy brain is labile and stable, able to adapt to new environments while pulling on knowledge and experience to make sense of the sensory world. Thus, in addition to motivating and informing interventions, our framework can help facilitate training by identifying an individual's strengths and weaknesses in the neural processes important for everyday communication.

Acknowledgments

The authors are grateful to Trent Nicol for his input on the manuscript and to colleagues in the Auditory Neuroscience Laboratory, past and present, for the many experiments and conversations that motivated this framework. Supported by the NIH (R01 HD069414), the NSF (BCS 1430400), and the Knowles Hearing Center.

References

- Webster, D.B. (1992) An overview of mammalian auditory pathways with an emphasis on humans. In *The Mammalian Auditory Pathway: Neuroanatomy*. pp. 1–22, Springer
- Peretz, I. and Coltheart, M. (2003) Modularity of music processing. *Nat. Neurosci.* 6, 688–691
- Winer, J.A. and Schreiner, C.E. (2005) The central auditory system: a functional analysis. In *The Inferior Colliculus*. pp. 1–68, Springer
- Saalmann, Y.B. et al. (2012) The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* 337, 753–756
- Bajo, V.M. and King, A.J. (2012) Cortical modulation of auditory processing in the midbrain. *Front. Neural Circuits* 6, 114
- Behrmann, M. and Plaut, D.C. (2013) Distributed circuits, not circumscribed centers, mediate visual recognition. *Trends Cogn. Sci.* 17, 210–219
- Atiani, S. et al. (2014) Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 82, 486–499
- Siegel, M. et al. (2015) Cortical information flow during flexible sensorimotor decisions. *Science* 348, 1352–1355
- Gordon, K.A. et al. (2003) Activity-dependent developmental plasticity of the auditory brain stem in children who use cochlear implants. *Ear Hear.* 24, 485–500
- Benasich, A.A. et al. (2014) Plasticity in developing brain: active auditory exposure impacts prelinguistic acoustic mapping. *J. Neurosci.* 34, 13349–13363
- Tierney, A. et al. (2015) Music training alters the course of adolescent auditory development. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10062–10067
- Kilgard, M.P. (2012) Harnessing plasticity to understand learning and treat disease. *Trends Neurosci.* 35, 715–722
- Skoe, E. and Kraus, N. (2012) A little goes a long way: how the adult brain is shaped by musical training in childhood. *J. Neurosci.* 32, 11507–11510
- White-Schwach, T. et al. (2013) Older adults benefit from music training early in life: biological evidence for long-term training-driven plasticity. *J. Neurosci.* 33, 17667–17674
- Anderson, S. et al. (2013) A dynamic auditory–cognitive system supports speech-in-noise perception in older adults. *Hear. Res.* 300, 18–32
- Shinn-Cunningham, B.G. (2008) Object-based auditory and visual attention. *Trends Cogn. Sci.* 12, 182–186
- Rönnberg, J. et al. (2008) Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47, S99–S105
- Nahum, M. et al. (2008) Low-level information and high-level perception: the case of speech in noise. *PLoS Biol.* 6, e126
- Kraus, N. and Chandrasekaran, B. (2010) Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11, 599–605
- Polley, D.B. et al. (2006) Perceptual learning directs auditory cortical map reorganization through top-down influences. *J. Neurosci.* 26, 4970–4982
- Kraus, N. et al. (2012) Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. *Ann. N. Y. Acad. Sci.* 1252, 100–107
- Green, C.S. and Bavelier, D. (2003) Action video game modifies visual selective attention. *Nature* 423, 534–537
- Anderson, S. et al. (2013) Reversal of age-related neural timing delays with training. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4357–4362
- Anguera, J. et al. (2013) Video game training enhances cognitive control in older adults. *Nature* 501, 97–101
- Patel, A.D. (2011) Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front. Psychol.* 2, 142
- Herholz, S.C. and Zatorre, R.J. (2012) Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76, 486–502
- Strait, D.L. and Kraus, N. (2014) Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. *Hear. Res.* 308, 109–121
- Limb, C.J. and Braun, A.R. (2008) Neural substrates of spontaneous musical performance: an fMRI study of jazz improvisation. *PLoS ONE* 3, e1679
- Blood, A.J. et al. (1999) Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* 2, 382–387
- Parbery-Clark, A. et al. (2009) Musical experience limits the degradative effects of background noise on the neural processing of sound. *J. Neurosci.* 29, 14100–14107
- Krizman, J. et al. (2012) Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7877–7881
- Swaminathan, J. et al. (2015) Musical training, individual differences and the cocktail party problem. *Sci. Rep.* 5, 11628
- Zendel, B.R. and Alain, C. (2012) Musicians experience less age-related decline in central auditory processing. *Psychol. Aging* 27, 410
- Boebinger, D. et al. (2015) Musicians and non-musicians are equally adept at perceiving masked speech. *J. Acoust. Soc. Am.* 137, 378–387
- Rogers, C.L. et al. (2006) Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Appl. Psycholinguist.* 27, 465–485

36. Van Engen, K.J. (2010) Similarity and familiarity: second language sentence recognition in first- and second-language multi-talker babble. *Speech Commun.* 52, 943–953
37. Strait, D.L. *et al.* (2009) Musical experience and neural efficiency – effects of training on subcortical processing of vocal expressions of emotion. *Eur. J. Neurosci.* 29, 661–668
38. Kraus, N. and Nicol, T. (2014) The cognitive auditory system. In *Perspectives on Auditory Research* (Fay, R. and Popper, A., eds), pp. 299–319, Springer-Verlag
39. Lin, F.R. *et al.* (2013) Hearing loss and cognitive decline in older adults. *JAMA Intern. Med.* 173, 293–299
40. Anderson, S. *et al.* (2012) Aging affects neural precision of speech encoding. *J. Neurosci.* 32, 14156–14164
41. Mishra, J. *et al.* (2014) Adaptive training diminishes distractibility in aging across species. *Neuron* 84, 1091–1103
42. Merzenich, M.M. *et al.* (1984) Somatosensory cortical map changes following digit amputation in adult monkeys. *J. Comp. Neurol.* 224, 591–605
43. Neville, H.J. *et al.* (1998) Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience. *Proc. Natl. Acad. Sci. U.S.A.* 95, 922–929
44. Micheyl, C. *et al.* (1997) Difference in cochlear efferent activity between musicians and non-musicians. *Neuroreport* 8, 1047–1050
45. Bidelman, G.M. *et al.* (2014) Psychophysical auditory filter estimates reveal sharper cochlear tuning in musicians. *J. Acoust. Soc. Am.* 136, EL33–EL39
46. Kraus, N. *et al.* (2014) Auditory learning through active engagement with sound: biological impact of community music lessons in at-risk children. *Front. Neurosci.* 8, 351
47. Engineer, N.D. *et al.* (2011) Reversing pathological neural activity using targeted plasticity. *Nature* 470, 101–104
48. Brainard, M.S. and Doupe, A.J. (2000) Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* 404, 762–766
49. Thomson, J.M. and Goswami, U. (2008) Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. *J. Physiol. Paris* 102, 120–129
50. Woodruff Carr, K. *et al.* (2014) Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14559–14564
51. Bhide, A. *et al.* (2013) A rhythmic musical intervention for poor readers: a comparison of efficacy with a letter-based intervention. *Mind Brain Educ.* 7, 113–123
52. Saffran, J.R. *et al.* (1996) Statistical learning by 8-month-old infants. *Science* 274, 1926–1928
53. Lev-Williams, C. and Saffran, J.R. (2012) All words are not created equal: expectations about word length guide infant statistical learning. *Cognition* 122, 241–246
54. Skoe, E. *et al.* (2014) Prior experience biases subcortical sensitivity to sound patterns. *J. Cogn. Neurosci.* 27, 124–140
55. Bakin, J.S. and Weinberger, N.M. (1996) Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. *Proc. Natl. Acad. Sci. U.S.A.* 93, 11219–11224
56. Kilgard, M.P. and Merzenich, M.M. (1998) Cortical map reorganization enabled by nucleus basalis activity. *Science* 279, 1714–1718
57. Leaver, A.M. *et al.* (2011) Dysregulation of limbic and auditory networks in tinnitus. *Neuron* 69, 33–43
58. De Ridder, D. *et al.* (2015) Placebo-controlled vagus nerve stimulation paired with tones in a patient with refractory tinnitus: a case report. *Otol. Neurotol.* 36, 575–580
59. Salimpoor, V.N. *et al.* (2013) Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science* 340, 216–219
60. Kuhl, P.K. (2007) Is speech learning “gated” by the social brain? *Dev. Sci.* 10, 110–120
61. Kuhl, P.K. *et al.* (2003) Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9096–9101
62. Abrams, D.A. *et al.* (2013) Underconnectivity between voice-selective cortex and reward circuitry in children with autism. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12060–12065
63. Russo, N. *et al.* (2008) Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clin. Neurophysiol.* 119, 1720–1731
64. Hart, B. and Risley, T.R. (1995) *Meaningful Differences in the Everyday Experience of Young American Children*, Paul H. Brookes
65. Skoe, E. *et al.* (2013) The impoverished brain: disparities in maternal education affect the neural response to sound. *J. Neurosci.* 33, 17221–17231
66. Zhu, X. *et al.* (2014) Environmental acoustic enrichment promotes recovery from developmentally degraded auditory cortical processing. *J. Neurosci.* 34, 5406–5415
67. David, S.V. *et al.* (2012) Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2144–2149
68. Xiong, Q. *et al.* (2015) Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* 521, 348–351
69. Kral, A. and Eggermont, J.J. (2007) What's to lose and what's to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Res. Rev.* 56, 259–269
70. Bajo, V.M. *et al.* (2009) The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat. Neurosci.* 13, 253–260
71. Ayala, Y.A. *et al.* (2015) Differences in the strength of cortical and brainstem inputs to SSA and non-SSA neurons in the inferior colliculus. *Sci. Rep.* 5, 10383
72. Krishnan, A. *et al.* (2005) Encoding of pitch in the human brainstem is sensitive to language experience. *Cogn. Brain Res.* 25, 161–168
73. Wright, B.A. *et al.* (2010) Enhancing perceptual learning by combining practice with periods of additional sensory stimulation. *J. Neurosci.* 30, 12868–12877
74. Sarro, E.C. and Sanes, D.H. (2011) The cost and benefit of juvenile training on adult perceptual skill. *J. Neurosci.* 31, 5383–5391
75. Green, C. and Bavelier, D. (2012) Learning, attentional control, and action video games. *Curr. Biol.* 22, R197–R206
76. Gao, E. and Suga, N. (2000) Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: role of the corticofugal system. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8081
77. Bajo, V.M. *et al.* (2007) The ferret auditory cortex: descending projections to the inferior colliculus. *Cereb. Cortex* 17, 475–491
78. Patel, A.D. (2014) The evolutionary biology of musical rhythm: was Darwin wrong? *PLoS Biol.* 12, e1001821
79. White-Schwoch, T. *et al.* (2015) Auditory processing in noise: a preschool biomarker for literacy. *PLoS Biol.* 13, e1002196
80. Schreiner, C.E. and Urbas, J.V. (1988) Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hear. Res.* 32, 49–63
81. Rauschecker, J.P. *et al.* (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268, 111–114
82. Zatorre, R.J. *et al.* (2002) Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46
83. Overath, T. *et al.* (2015) The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nat. Neurosci.* 18, 903–911
84. Kraus, N. *et al.* (1988) Midline and temporal lobe MLRs in the guinea pig originate from different generator systems: a conceptual framework for new and existing data. *Electroencephalogr. Clin. Neurophysiol.* 70, 541–558
85. Kraus, N. *et al.* (1994) Discrimination of speech-like contrasts in the auditory thalamus and cortex. *J. Acoust. Soc. Am.* 96, 2758–2768
86. Abrams, D.A. *et al.* (2011) A possible role for a paralemniscal auditory pathway in the coding of slow temporal information. *Hear. Res.* 272, 125–134
87. Ahissar, E. *et al.* (2000) Transformation from temporal to rate coding in a somatosensory thalamocortical pathway. *Nature* 406, 302–306

88. Merigan, W.H. and Maunsell, J.H. (1993) How parallel are the primate visual pathways? *Annu. Rev. Neurosci.* 16, 369–402
89. Kraus, N. *et al.* (1996) Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science* 273, 971–973
90. Abrams, D.A. *et al.* (2010) Rapid acoustic processing in the auditory brainstem is not related to cortical asymmetry for the syllable rate of speech. *Clin. Neurophysiol.* 121, 1343–1350
91. Goswami, U. (2011) A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 15, 3–10
92. Kraus, N. *et al.* (1994) Nonprimary auditory thalamic representation of acoustic change. *J. Neurophysiol.* 72, 1270–1277
93. Kraus, N. and Disterhoft, J.F. (1982) Response plasticity of single neurons in rabbit auditory association cortex during tone-signal learning. *Brain Res.* 246, 205–215
94. Tallal, P. and Piercy, M. (1973) Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241, 468–469
95. Galbraith, G.C. *et al.* (1995) Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport* 6, 2363–2367
96. Cunningham, J. *et al.* (2001) Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. *Clin. Neurophysiol.* 112, 758–767
97. Anderson, S. *et al.* (2013) Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance. *J. Speech Lang. Hear. Res.* 56, 31–43
98. Rocha-Muniz, C.N. *et al.* (2014) Sensitivity, specificity and efficiency of speech-evoked ABR. *Hear. Res.* 28, 15–22
99. Tarasenko, M.A. *et al.* (2014) The auditory brainstem response to complex sounds: a potential biomarker for guiding treatment of psychosis. *Front. Psychiatry* 5, 142
100. White-Schwoch, T. *et al.* (2015) Auditory-neurophysiological responses to speech during early childhood: effects of background noise. *Hear. Res.* 328, 34–47
101. Anderson, S. *et al.* (2015) Development of subcortical speech representation in human infants. *J. Acoust. Soc. Am.* 137, 3346–3355
102. Skoe, E. *et al.* (2015) Stability and plasticity of auditory brainstem function across the lifespan. *Cereb. Cortex* 25, 1415–1426
103. Warrier, C.M. *et al.* (2011) Inferior colliculus contributions to phase encoding of stop consonants in an animal model. *Hear. Res.* 282, 108–118
104. Vélez, A. *et al.* (2013) Anuran acoustic signal perception in noisy environments. In *Animal Communication and Noise*. pp. 133–185, Springer
105. Krause, J. *et al.* (2007) The derived *FOXP2* variant of modern humans was shared with Neandertals. *Curr. Biol.* 17, 1908–1912
106. van Rheede, J.J. *et al.* (2015) Sensory-evoked spiking behavior emerges via an experience-dependent plasticity mechanism. *Neuron* 87, 1050–1062
107. Recanzone, G.H. *et al.* (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* 13, 87–103
108. Robertson, D. and Irvine, D.R.F. (1989) Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J. Comp. Neurol.* 282, 456–471
109. Edeline, J.-M. *et al.* (2011) Induction of selective plasticity in the frequency tuning of auditory cortex and auditory thalamus neurons by locus coeruleus stimulation. *Hear. Res.* 274, 75–84
110. Froemke, R.C. *et al.* (2013) Long-term modification of cortical synapses improves sensory perception. *Nat. Neurosci.* 16, 79–88