Chapter 6
The Janus Face of Auditory Learning: How Life in Sound Shapes Everyday Communication

Travis White-Schwoch and Nina Kraus

Abstract Human communication is rooted in sound. The frequency-following response (FFR) reveals the integrity of sound processing in the brain and characterizes sound processing skills in a diverse group of listeners spanning a gamut from expertise to disorder. These neurophysiological processes are shaped by life experiences, for better or worse. Here two lifelong experiences are juxtaposed—music training (a model of expertise) and low socioeconomic status (a model of disorder)—with an emphasis on how FFR studies reveal their combined influences on everyday communication skills. A view emerges wherein the auditory system is an integrated, but distributed, processing circuit that interacts with other modalities, cognitive systems, and the limbic system to shape auditory physiology. Objective indices of auditory physiology, such as the FFR, therefore can be thought of as measures of "whole brain" auditory processing that recapitulate the past by revealing the legacy of prior experiences and predict the future by elucidating their consequences for everyday communication.

T. White-Schwoch
Auditory Neuroscience Laboratory, Department of Communication Sciences,
Northwestern University, 2240 Campus Dr., Evanston, IL 60208, USA
e-mail: whiteschwoch@northwestern.edu

N. Kraus (✉)
Auditory Neuroscience Laboratory, Departments of Communication Sciences,
Neurobiology and Otolaryngology, Northwestern University,
2240 Campus Dr., Evanston, IL 60208, USA
e-mail: nkraus@northwestern.edu

6.1   Introduction

Janus, the Roman god of beginnings, transitions, and endings, is classically depicted with two faces: one that looks forward and one that looks back (Fig. 6.1). In this regard he symbolizes two interrelated principles of auditory learning that are manifest in the frequency-following response (FFR). First, auditory learning is inherently a double-edged sword that shapes neural function for better or for worse. Second, auditory physiology recapitulates the past by reflecting the legacy of these experiences and predicts the future by delineating their consequences for everyday communication. Thus, a single FFR is a snapshot of an individual's auditory processing past and potential.

These two principles are reviewed in this chapter with an emphasis on the enduring biological legacy imparted by life in sound and the presumed consequences of auditory experiences for everyday communication. It is important to keep in mind that these principles apply to the auditory system in general; the FFR simply provides an experimental avenue to explore them in humans and to relate these findings to animal models.

Here, FFR studies of everyday communication skills and how these skills are shaped by experience are reviewed. Two long-term experiences are contrasted: music, a case of enrichment, and low socioeconomic status (SES), a background associated with disorders of brain and cognition functions. Juxtaposing these discrepant experiences illustrates how overlapping circuitry is involved in states of expertise and disorder and suggests they arise through similar mechanisms. Importantly, the idea that similar, malleable neural pathways are involved in states

Fig. 6.1 The Janus face of auditory learning. Janus, the Roman god of beginnings, transitions, and endings, is classically depicted with two faces. His two faces symbolize two principles of auditory learning. First, learning is a double-edged sword, which can shape neural function for better or worse. Second, auditory physiology recapitulates the past by reflecting the biological legacy of a life in sound and predicts the future by delineating the consequences of those experiences for everyday communication. (Photograph from www.wikipedia.org)

of expertise and disorder motivates experience-dependent interventions to improve listening (Carcagno and Plack, Chap. 4). This idea cuts across clinical populations, including individuals with auditory processing disorder (Schochat, Rocha-Muniz, and Filippini, Chap. 9), individuals with reading impairment (Reetzke, Xie, and Chandrasekaran, Chap. 10), and older adults (Anderson, Chap. 11).

6.2 The Auditory System: Distributed but Integrated

Classic models of the auditory system emphasize its hierarchical nature, that is, how each station along the neuraxis is specialized (Webster 1992). Despite the theoretical contributions of these models to auditory neuroscience, they risk failing to account for the remarkable integration the auditory system achieves. After all, sound catalyzes activity *throughout* the auditory neuraxis, from cochlea to cortex and back again. A complete understanding of how the brain makes sense of sound therefore has to consider not just what each auditory relay uniquely contributes to auditory processing, but how they integrate (Kraus and White-Schwoch 2015). This integration occurs along with complementary brain networks, such as cognitive and limbic circuits.

Thus, a view emerges wherein the auditory system is *distributed but integrated*. Any measure of auditory physiology, including the FFR, can be thought to index "whole brain" auditory processing, insofar as it reflects activity that occurs in concert with the rest of the system and that has been shaped in the context of that system. This view does not ignore the fact that different measures of auditory function have different generators: otoacoustic emissions (OAEs) are generated by the outer hair cells, and the FFR is generated primarily by the auditory midbrain. But both hair cell and midbrain physiology are shaped by experience, meaning OAEs and FFRs reflect this experience. This network-based account of auditory physiology relates to emerging models of sensory systems that emphasize the "push and pull" among sensory organs, subcortical networks, and cortical networks (Behrmann and Plaut 2013; Pafundo et al. 2016).

The FFR encapsulates this view. As a simple case in point, consider the extraordinary diversity of research conducted with the FFR, spanning lifelong language experience (Krishnan and Gandour, Chap. 3) to in-the-moment adaptation (Escera, Chap. 5). Consider also the broad clinical applications studied with the FFR for auditory processing disorders (Schochat, Rocha-Muniz, and Filippini, Chap. 9), reading impairment (Reetzke, Xie, and Chandrasekaran, Chap. 10), and age-related hearing impairment (Anderson, Chap. 11). Needless to say, these three clinical populations do not share a common lesion in the auditory midbrain; rather, the midbrain functions in the context of bottom-up and top-down networks that shape its response properties, making the FFR a viable tool in the laboratory and clinic. This diversity of research employing the FFR highlights how it is a measure of integrative auditory function (Coffey et al. 2016).

6.3   Conceptual Frameworks for Auditory Learning

The premise that motivates studies of auditory experience is that sound provides a fundamental medium for learning about the world. Thus, life in sound (particularly during early childhood) shapes both auditory and nonauditory social and cognitive brain processes.

6.3.1   Maturation by Experience

The *maturation-by-experience hypothesis* posits that sensory experience is necessary for the development of sensory circuitry. This idea has been studied extensively, especially with respect to language development. For example, Kuhl et al. (1992) have shown that newborns are sensitive to sound contrasts that are not meaningful in their native language, but this sensitivity narrows by approximately six months-of-age, suggesting that ongoing exposure to the linguistic environment shapes basic perception. These studies are complemented by physiological studies that show how language experiences shape sound processing (Jeng, Chap. 2; Krishnan and Gandour, Chap. 3).

Additional evidence for maturation by experience comes from studies of cochlear implantation. Gordon et al. (2012) have shown that the auditory brainstem matures as a function of auditory experience by measuring electrically evoked auditory brainstem responses (eABRs) in children with cochlear implants. They have documented a systematic relationship between eABR latency and years since implantation, suggesting that ongoing auditory experience modulates brainstem function.

The idea that sensory experience is tied to sensory development is also supported by evidence from the visual system. For example, van Rheede et al. (2015) show that visual input rapidly drives formerly dormant neurons in the tadpole optic tectum (a midbrain visual system structure) to spike in response to future visual input. In other words, visual experience creates the conditions for visual neurons to fire. Noteworthy is that van Rheede and colleagues show that experience is required for a fundamental unitary phenomenon—firing an action potential—thus highlighting how experience and perception are intrinsically intertwined.

The maturation-by-experience hypothesis motivates a strong prediction: if sensory experience is required for the development of sensory circuitry, then altering said experiences should alter the default state of said circuitry. In other words, two individuals reared in different sensory environments should exhibit distinct profiles in adulthood. This prediction is borne out by investigations of language experience (Intartaglia et al. 2016; Krishnan and Gandour, Chap. 3), music experience, and socioeconomic status. This prediction is also supported by studies in animals that have been raised in enriched environments and exhibited more sophisticated sensory processing (audition: Engineer et al. 2004; vision: Wang et al. 2010) compared

to those that have been raised in toxic environments and exhibited less mature processing (Chang and Merzenich 2003).

These investigations do not rule out the clear role that predispositions, such as genetics, play in shaping adult brain circuitry. Nevertheless, they highlight how predispositions interact with experience in childhood and adulthood to sculpt a sensory phenotype. It should also be emphasized that early experiences influence future learning and plasticity in addition to what becomes automatic sensory processing.

6.3.2   The Cognitive-Sensorimotor-Reward Framework for Auditory Learning

Kraus and White-Schwoch (2015) propose a *cognitive-sensorimotor-reward* framework for auditory learning. Cognitive factors optimize learning because listening is an active process, and repeated engagement of cognitive systems sculpts sensory infrastructure to process sound more effectively. Sensorimotor systems optimize learning through the convergence of processing stations within and across modalities, suggesting that sensory neurons can be thought of as a form of "memory storage" because their intrinsic response properties retune to reflect past experience. Limbic systems optimize learning via neuromodulatory input throughout the auditory neuraxis.

It should be mentioned that these three ingredients may not be necessary for learning. For example, listeners are sensitive to the statistics of a soundscape and the auditory system rapidly adapts to exploit these regularities (Escera, Chap. 5). The focus of the cognitive-sensorimotor-reward framework, however, is long-term experience that reshapes fundamental sensory infrastructure. Noteworthy is that these enduring changes are reflected by the FFR, including in studies of the effects of music experience and poverty (Sects. 6.5 and 6.6).

6.3.3   The Afferent/Efferent, Primary/Non-primary Trade-off Model

A major view in auditory neuroscience is that the mechanisms responsible for in-the-moment plasticity, such as attending to one sound among many, are also responsible for large-scale functional remodeling following prolonged and repeated experiences (Fritz et al. 2003; Weinberger 2004). The idea is that plastic experiences, such as a repeated evanescent change in a cell's receptive field, eventually are engrained in that cell's intrinsic response properties.

The *afferent/efferent, primary/non-primary trade-off model* (Kraus and White-Schwoch 2015) contends that in-the-moment changes are qualitatively

distinct from long-term changes to neural function. This view does not discount clear evidence that both selective attention and long-term experiences engender forms of neuroplasticity, nor does it discount evidence that cellular consequences of these experiences are similar. Rather, it distinguishes between an active *learning* phase and a *learned* phase (e.g., Reed et al. 2011).

Kraus and White-Schwoch (2015) propose this evanescent-enduring dichotomy is rooted in the parallel primary and non-primary streams of the auditory neuraxis. The primary (lemniscal) pathway consists of "core" stations (such as the central nucleus of the inferior colliculus, the ventral division of the medial geniculate, and the primary auditory cortex). The lemniscal pathway is strongly tonotopic and characterized by short-latency, stimulus-dependent, unitary responses with a bias to tonal stimuli. The non-primary pathway (paralemniscal) consists of "belt" or "parabelt" stations (such as the lateral cortex of the inferior colliculus, the shell nucleus of the medial geniculate, or peripheral auditory fields). The paralemniscal pathway is characterized by less tonotopic response profiles and connectivity, more context-dependent coding and, in human cortex, a tendency to specialize for certain sounds such as speech (reviewed in Abrams et al. 2011). Additionally, the primary pathway is biased to process faster sounds, whereas the non-primary pathway is biased to process slower sounds (Ahissar et al. 2000; Abrams et al. 2011). Finally, it has been speculated that these parallel pathways may underlie the remarkable balance struck between stability and flexibility in auditory processing (Kraus and White-Schwoch 2015).

Both primary and non-primary divisions of the auditory neuraxis contain afferent (ear-to-brain) and efferent (brain-to-ear) projections. Kraus and White-Schwoch (2015) propose a trade-off between the relative weights of these projections between the two streams (Fig. 6.2). Specifically, they argue that the primary stream is biased to afferent processing that reflects deeply engrained response profiles with relatively fewer resources dedicated to efferent processing. In contrast, they argue that the non-primary stream is biased to efferent processing that facilitates rapid task-dependent plasticity with relatively fewer resources dedicated to afferent processing.

This framework motivates the prediction that rapid task-dependent plasticity (i.e., the learning phase) should originate in non-primary auditory fields. Indeed, in ferret models, active attending drives spectrotemporal receptive field plasticity in posterior auditory fields prior to (and to a much greater degree than) in primary auditory cortex (Atiani et al. 2014). Moreover, recordings from rat auditory cortex show that stimulus-specific adaptation (SSA)—rapid plasticity occurring as a function of stimulus predictability (see Escera, Chap. 5)—occurs more quickly and robustly in non-primary auditory fields than in primary auditory cortex (Nieto-Diego and Malmierca 2016). This non-primary SSA is the presumed neural substrate of the mismatch negativity response (MMN, a scalp-recorded response reflecting detection of stimulus deviants). Similarly, recordings from guinea pigs show that non-primary auditory thalamus contributes to the MMN (Kraus et al. 1994a, b). Also noteworthy is that SSA-sensitive neurons in IC are strongly biased
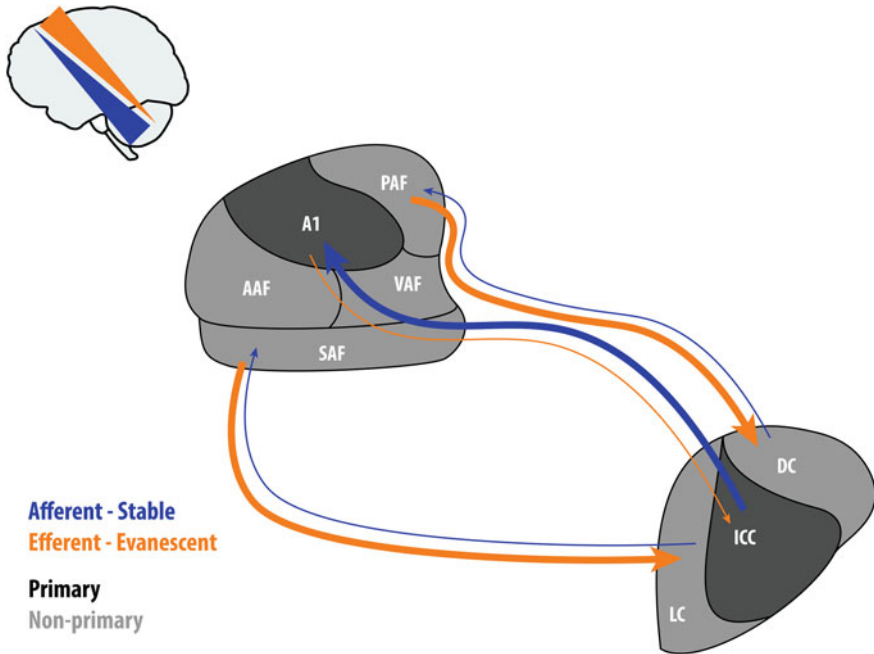
Fig. 6.2 The afferent/efferent, primary/non-primary trade-off model. The auditory neuraxis contains primary and non-primary divisions, each of which are innervated by afferent (*blue*) and efferent (*orange*) projections. Kraus and White-Schwoch (2015) propose a model wherein the primary division is biased to stable, afferent processing that is only remodeled over time. In contrast, the non-primary division is biased to efferent processing that is constantly undergoing evanescent plasticity. This idea is illustrated by schematized projections between the auditory cortex and inferior colliculus, showing primary (*darker*) and non-primary (*lighter*) fields, but applies throughout the system. (*A1*, primary auditory cortex; *AAF*, anterior auditory field; *DC*, dorsal cortex; *ICC*, central nucleus of inferior colliculus; *LC*, lateral cortex; *PAF*, posterior auditory field; *SAF*, superior auditory field; *VAF*, ventral auditory field. (Original figure by White-Schwoch and Kraus; auditory cortex view after Nieto-Diego and Malmierca 2016; inferior colliculus view after Loftus et al. 2008)

toward non-primary fields and receive input from non-primary corticocollicular fibers (Ayala et al. 2015).

This framework motivates a second prediction, namely, that primary auditory fields should only be resculpted following prolonged experiences (i.e., in the learned phase). This prediction is more difficult to test, especially given the challenges of following a single neuron over time. There is, however, evidence from the somatosensory system that cortical maps are reshaped following partial amputation (and restoration) of a finger (Merzenich et al. 1983). Studies in congenitally deaf humans and cats show that altered, daily sensory experience leads to drastic reorganization of primary auditory cortex (Kral and Sharma 2012). Finally, primary corticocollicular fibers are necessary for recalibrating sound localization abilities to

accommodate an earplug, a form of auditory learning (Bajo et al. 2010). Together, this shows that primary remodeling accumulates following intensive, prolonged experiences, such as a permanent alteration of sensory input. This idea is consistent with Kilgard's (2012) *expansion-renormalization model*, which argues that cortical map plasticity is a transient phenomenon, not necessary for changes in long-term performance.

Thus, it appears that rapid plasticity is qualitatively distinct from long-term functional remodeling. Evidence for this principle is found in several sites along the auditory neuraxis, reinforcing the review of a distributed-but-integrated circuit. With respect to the FFR, this model can explain a paradox that pervades the field. Although the FFR profoundly illustrates the legacy of prolonged and repeated experiences (see Sects. 6.5–6.7), evanescent changes due to attentional modulation are difficult to capture (see Shinn-Cunningham, Varghese, Wang, and Bharadwaj, Chap. 7). The FFR may be biased to index the more enduring and automatic processing route in the auditory system, which is only reshaped following experience. Noteworthy is the observation that online changes in inferior colliculus receptive fields are relatively small and in the opposite direction of auditory cortex receptive field plasticity induced by the same task (Slee and David 2015).

6.3.4 Selective Modulation of Sensory Function

An important principle of auditory neuroplasticity is that learning occurs along the dimensions that are trained, that is, the specific acoustic parameters that a listener attends to and connects to meaning are those whose neural coding is modulated. The *selective modulation* occurs in contrast to an overall "volume knob" effect that could be thought of as a broad gain (or decrease) in neural activity.

This principle was demonstrated elegantly in an experiment by Polley et al. (2006), who trained two groups of rats on a single stimulus set. One group trained on frequency contrasts between the stimuli, whereas another group trained on intensity contrasts. Perceptual acuity increased along the specific dimensions on which the rats were trained. Additionally, receptive fields in auditory cortex improved only along the trained dimension (frequency versus intensity). This experiment highlights how cognitive factors interact with sensory factors during learning. It also illustrates how plasticity is constrained to the factors that a listener is taught to care about as opposed to a product of the trained stimuli per se.

The FFR profoundly illustrates the principle of selective modulation of sensory function. A single response presents a plethora of information about how discrete components of sound are processed, and a central hypothesis behind much FFR research is that these reflect non-overlapping processes (see Kraus, Anderson, and White-Schwoch, Chap. 1). That is to say, discrete FFR parameters reflect distinct mechanisms that are modulated by divergent auditory experiences. This concept is broadly consistent with the idea that neuroplasticity occurs along dimensions that are behaviorally relevant to a listener (Recanzone et al. 1993). Additionally, it is

consistent with the idea of the auditory system as a distributed, but integrated, circuit insofar as a single index of neurophysiological function reveals multiple aspects of auditory processing.

6.3.5 Experimental Models of Auditory Learning in Humans

A major challenge in auditory neuroscience has been to bridge research in animal models and research on humans. Animal models offer clear advantages, including careful control over learning, environments, and genetics, in addition to the ability to make invasive recordings of neural activity. There are also several limitations with animal models, however, the largest being an inability to evaluate outcomes of learning for complex behaviors (such as communication) and the fact that laboratory conditions are rarely naturalistic. The FFR fills this gap by providing a robust index of auditory function across species (e.g., Warrier et al. 2011; White-Schwoch et al. in press).

Both music training and socioeconomic status (SES) provide models for auditory learning in humans because they both encapsulate a constellation of factors that drive neural remodeling. In other words, they provide "best case scenarios" for considering what is possible in terms of neuroplasticity. Music training provides a model of *adaptive plasticity*—gains in perceptual and cognitive skills exhibited in the learned phase. In contrast, SES provides a model of *maladaptive plasticity*—poorer perceptual and cognitive skills in the learned phase. Additional models include language experience (Krishnan and Gandour, Chap. 3) and perceptual learning (Carcagno and Plack, Chap. 4). An understanding of this learning, however, needs to be grounded in an understanding of everyday communication, its impairments, and their relationships to the FFR.

6.4 FFR Studies of Everyday Communication

FFR studies provide a foundation for informing the biology of auditory experience by elucidating not just the neural remodeling engendered by these experiences but the consequences of this remodeling for communication. This literature specifically lays the groundwork for understanding FFR studies of music training and SES. On the one hand, seeing that the same neurophysiological processes implicated in communication disorders are boosted by music training makes a strong case for therapeutic interventions. On the other hand, seeing that the neurophysiological profile of communication disorders partially overlaps that of SES can help understand the linguistic and cognitive sequelae of growing up in low SES environments. Finally, showing how FFR parameters do—and do not—overlap between different populations reinforces the idea that plasticity is a double-edged sword that occurs

along dimensions of behavioral relevance. Thus, an understanding of FFR studies of everyday communication provides a context for studies of long-term auditory experience by highlighting their consequences for everyday behavior.

6.4.1 Listening and Language Are Connected

FFR studies of communication skills are rooted in the idea that listening, language, and literacy are connected. Many models of communication impairments, including reading impairment and autism spectrum disorder, identify abnormal auditory processing as a core deficit (e.g., Tallal 2004). A child who cannot efficiently connect sounds to meaning will create a poor infrastructure on which to develop skills such as reading. Although it remains debated whether this abnormal auditory perception is a cause or consequence of poor language skills (Rosen 2003), objective measures (such as the FFR) prove advantageous in evaluating communication skills in these populations (Reetzke, Xie, and Chandrasekaran, Chap. 10). Specifically, because the FFR does not require a behavioral response from listeners, it may be applied to difficult-to-test populations such as infants (see Jeng, Chap. 2) or children with attention problems (see Schochat, Rocha-Muniz, and Filippini, Chap. 9).

The FFR has clarified that listening skills (such as understanding speech in noisy environments) and language skills (such as reading) only partially overlap. Listening in noise pulls on the ability to group talkers and auditory objects, whereas reading pulls on the ability to categorize sounds into phonemes. Both rely on accurate transcription of speech sounds and sensitivity to the auditory environment. Thus, individuals with communication impairments can struggle in one or both domains. Importantly, FFR studies of auditory enrichment and deprivation have also shown that the same facets of sound processing important for communication are malleable through experience (Kraus and White-Schwoch 2016), thereby motivating therapeutic approaches to improve the neural foundations of everyday communication.

6.4.2 Everyday Communication Skills are Revealed by the FFR

6.4.2.1 Phase One: FFRs are Distinct in Children with Learning Problems

The first set of FFR studies of communication skills involved a broadly defined group of children with auditory-based learning problems (LPs) such as dyslexia, specific language impairment (SLI), auditory processing disorder (APD), and

attention-deficit/hyperactivity disorder. The emphasis of this set of studies was to establish group differences between LP children and their typically developing peers. Together, these studies provide a proof of concept that the FFR characterizes auditory-neurophysiological processing in children with and without communication impairments.

Cunningham et al. (2001) conducted the first FFR study of LP children. They compared FFRs to the consonant-vowel (CV) syllable [da] in quiet and background noise. The LP children had slower responses than their peers and less robust representation of the speech harmonics that conveyed speech formant features. When the FFRs were correlated to the evoking stimulus, the LP children had smaller scores, suggesting a less faithful representation of sound. However, group differences were only apparent when the [da] was played in noise. Consistent with the integrated-but-distributed model of the auditory system (see Sect. 6.2), these children also had smaller cortical responses to speech in noise. Finally, they required larger acoustic differences to distinguish two similar syllables presented in noise.

King et al. (2002) followed up on this work and replicated the finding that LP children have slower responses to CV syllables concomitant to less robust cortical responses to speech. Wible et al. (2004) showed LP children had less-refined onset responses to speech, suggesting a loss of neural synchrony, especially for rapidly presented streams of stimuli. Again, consistent with the distributed-but-integrated view of the auditory system, Abrams et al. (2006) found that FFRs to speech relate to the integrity with which rapid temporal cues are processed in auditory cortex; however, this coherence is disrupted in LP children (Wible et al. 2005).

Together, these studies show that children with communication problems broadly spanning language, literacy, and auditory processing, have poorer speech-evoked FFRs than their peers. Specifically, their FFRs tend to be slower, smaller, and to reproduce stimulus features less accurately. When sounds are presented in noise or rapid succession, group differences become more apparent, suggesting that listening conditions that tax the system emphasize neurophysiological processing bottlenecks in LP children.

6.4.2.2 Phase Two: FFRs Reveal Mechanistic Bottlenecks in Sound Processing

The second phase aimed to elucidate direct relationships between FFR parameters and communication behaviors such as literacy skills or the ability to understand speech in noise. This phase is noteworthy for introducing more diverse stimuli and complex paradigms and techniques. This work reinforces the relationship between the ability to understand speech in everyday environments and literacy skills and points to potential auditory-neurophysiological mechanisms underlying this link. Finally, this phase of work introduced the idea of partially overlapping "neural signatures" for communication—patterns of enhancements and diminutions to sound processing manifest by the FFR that characterize groups of listeners (Kraus and Nicol 2014).

Banai et al. (2009) investigated phonological processing (knowledge of the sound structure of spoken language; see Reetzke, Xie, and Chandrasekaran, Chap. 10) in a group of children with a wide range of reading skills. Children with poor phonological processing had slower FFRs than their peers with better phonological processing. Additionally, these children had smaller responses to speech harmonics. Notably, Banai et al. (2009) showed that groups of children had similar responses to the fundamental frequency (F_0), which is consistent with the idea of selective modulation of neurophysiological functions (see Sect. 6.3.4). Thus, children with poor phonological skills have poor coding of speech formant features but intact coding of the F_0 (Kraus and Nicol 2005).

Chandrasekaran et al. (2009) introduced a new paradigm to measure FFRs: instead of eliciting responses to a single stimulus, they presented eight CV syllables in a random order, one of which was also presented in a continuous stream in a second recording session. The hypothesis was that the auditory system would be sensitive to stimulus context, motivated by evidence that neural coding is modulated by the statistics of sounds in the environment (see Escera, Chap. 5). Indeed, Chandrasekaran et al. (2009) found a stimulus adaptation effect—responses to harmonics in the stimulus were amplified when it was played in a predictable condition. Interestingly, the extent of this amplification was related to children's ability to understand sentences in noise, and it was attenuated in children with dyslexia. This finding ties into Ahissar's (2007) argument that a primary deficit in dyslexia is an inability to exploit regularities in a sensory stream.

Hornickel et al. (2009) measured FFRs to the contrastive CV syllables [ba], [da], and [ga]. The three stimuli are expected to elicit a stereotypical response pattern based on their acoustics (see Fig. 6.3). The extent of FFR timing differences related both to the ability to understand sentences in noise and to phonological processing. Additionally, poor readers did not exhibit the expected timing pattern between responses to the three sounds, suggesting they have "blurry" neurophysiological encoding of these speech features (Fig. 6.3).

Together, these studies reinforce the idea that accurate representation of fine-grained speech features is crucial for literacy and language development and the idea that everyday listening skills support reading development (Tallal 2004).

Hornickel and Kraus (2013) investigated a different aspect of neural coding with respect to literacy. Instead of asking how quickly or accurately stimulus features are coded, they quantified the across-trial variability of the FFR. They found a systematic relationship between reading fluency and FFR variability: the best readers had the most consistent responses to speech and the poorest readers had the most variable responses. This suggests that children with poor literacy skills do not process sensory input reliably, which presumably hampers the development of a robust knowledge of language. Noteworthy is a similar phenomenon in auditory cortex of rats with a dyslexia candidate gene knockdown (Centanni et al. 2014).

Anderson et al. (2010a) focused on children's abilities to understand sentences in noise. They compared FFRs to [da] presented in quiet and background noise and quantified the extent to which noise delayed timing in response to the consonant-vowel transition. Children with better sentence perception in noise
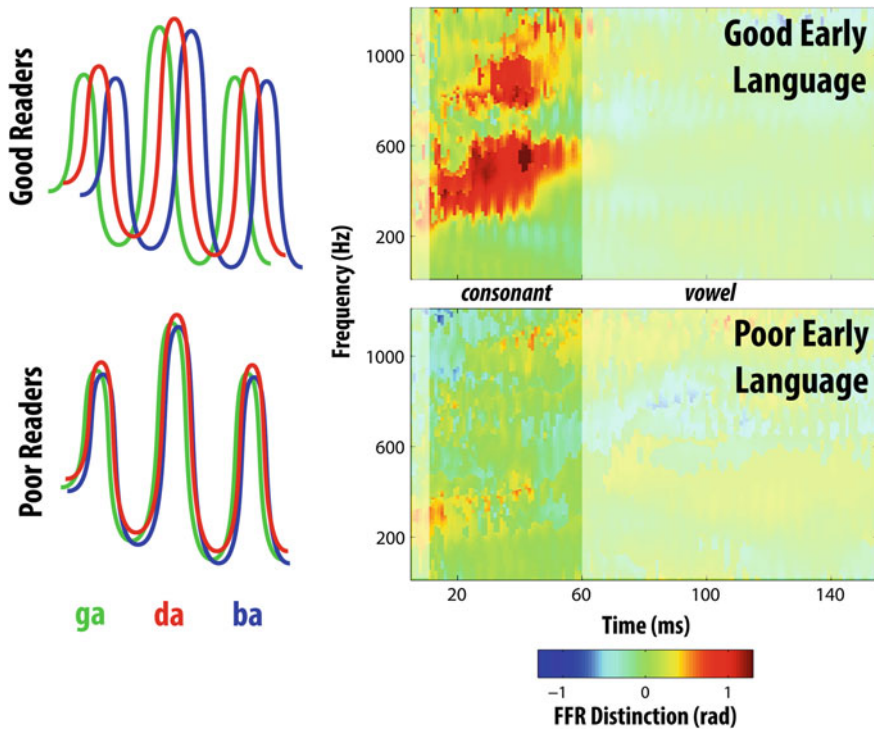
Fig. 6.3 FFR distinction of contrastive sounds. *Left* FFRs to the sounds [ba], [da], and [ga] are expected to have different timing based on the acoustics of the sounds. /G/ is the highest frequency sound, and so should elicit the earliest response, followed by /d/, followed by /b/. Good readers show this pattern, but poor readers do not. (Adapted from Hornickel et al. 2009). *Right* The cross-phaseogram measures the extent to which these sounds are distinguished in time-frequency space. Stronger distinctions (deeper red colors) are found in preschoolers with better early language skills. (Adapted from White-Schwoch and Kraus 2013)

exhibited a smaller FFR timing delay. Additionally, children with stronger reading fluency (the ability to quickly and accurately parse text) exhibited a smaller timing delay in noise. This study reinforced the link between FFR timing and communication skills and the link between literacy and perception in noise. In a follow-up study, however, they showed that children with better sentence perception in noise had stronger representation of the F_0 in speech (Anderson et al. 2010b). Thus, while the neural signatures for literacy and hearing in noise seemed to overlap with respect to neural response timing, they differed in that hearing in noise was associated with representation of the F_0, whereas literacy was associated with representation of speech harmonics.

The link between F_0 coding and the ability to understand sentences in noise was also observed in young adults whose F_0 coding and sentence-in-noise perception are both strengthened following auditory training (Song et al. 2012). This F_0

hearing-in-noise contingency has also been found in older adults with normal hearing; thus, F_0 coding appears to underlie listening in noise across the life span (see Bidelman, Chap. 8 and Anderson, Chap. 11).

Basu et al. (2010) focused on children with specific language impairment (SLI), a disorder that affects language processing and production, including phonics, grammar, and pragmatics (Bishop 1997). Many children with SLI exhibit abnormal perception of rapidly changing sounds (e.g. Benasich and Tallal 2002). Children with SLI had poorer or absent phase locking in response to rapidly changing, high-frequency sweeps, and smaller spectral amplitudes in response to fast-moving frequency sweeps (see Krishnan and Gandour, Chap. 3 for a discussion of FFR phase locking to tonal sweeps). Additionally, children with SLI exhibited a longer timing delay when presentation rate increased, consistent with the idea that taxing the system pulls out differences in clinical populations (Sect. 6.4.2.1).

Russo et al. (2008) focused on high-functioning children with autism spectrum disorder (ASD), a wide-ranging disorder of language and communication that affects social and pragmatic cues in language. They measured FFRs to the sound [ya] with two prosodic contours: one that rose in pitch (as in a question) and one that fell in pitch (as in a declaration). A majority of children with ASD did not accurately track the pitch of the sound in their FFRs. This provides neurophysiological evidence that the auditory coding of prosodic cues in speech, which often convey emotion, humor, and intention, is disrupted in children with ASD. Additionally, children with ASD had slower responses to the onset and formant features of the CV syllable [da] (Russo et al. 2009).

Together, these studies show how FFR signatures evoke the hallmark behavioral phenotype of the communication disorder. In other words, the coding of specific acoustic and acoustic-phonetic parameters that challenge distinct groups of listeners is disrupted (e.g., formant transitions in dyslexia, prosody in ASD, etc.). Additionally, this literature illustrates how neural signatures of distinct communication abilities and disabilities only partially overlap. If different communication impairments have distinct neural signatures, then the FFR may be a viable tool for evaluating auditory processing in clinical populations (Anderson and Kraus 2016).

6.4.2.3 Phase Three: FFRs Predict the Future

Recent attention in FFR studies has turned to the potential for the FFR to predict the development of communication skills. As reviewed by Jeng (Chap. 2), FFRs are robust during infancy and in preschoolers (also see Anderson et al. 2015). The long-term goal of this line of research is therefore to identify FFR features that predict communication impairments before children struggle with language development. Paradoxically, many disorders cannot be identified until a child has struggled and failed. For example, a diagnosis of dyslexia requires that a child has received prolonged formal instruction, setting opportunities for remediation back several years (Ozernov-Palchik and Gaab 2016). This goal is motivated by the intuitive fact that early interventions are extremely effective. Bishop and Adams (1990) show that if a

child's language problems are resolved by age 5.5 years, literacy development pro-
ceeds smoothly; otherwise, literacy challenges may be anticipated. This is not to say
that interventions later in life are ineffective (see Carcagno and Plack, Chap. 4); rather,
because they cannot piggyback a sensitive period for auditory development (see
Sect. 6.3.1) interventions likely have to be more intense (but see Sect. 6.7.1.2).

White-Schwoch and Kraus (2013) replicated Hornickel and colleagues' (2009)
study of speech sound distinction in the FFR (see Sect. 6.4.2.2) in preschool
children who had not yet learned to read. They found that pre-readers with stronger
early phonological processing had stronger neurophysiological distinctions of
contrastive sounds (Fig. 6.3). This shows a neural correlate of a core literacy skill
that precedes formal instruction in reading, suggesting poor neurophysiological
processing is a bottleneck in literacy development (as opposed to a consequence of
limited reading experience). Additionally, this suggests that preschool FFRs iden-
tify children with early constraints on literacy development.

Woodruff Carr et al. (2014) conducted a study in a similar population focused on
rhythm skills, motivated by evidence that individuals with dyslexia struggle to
maintain a steady beat (Goswami 2011). Preschoolers who could maintain a steady
beat had more accurate FFR coding of the temporal envelope of a speech syllable.
Consistent with the hypothesis that sensitivity to envelope cues is crucial for lit-
eracy development (Goswami 2011), these children also had more advanced early
literacy skills than their peers. Thus, it once again seems that neural correlates of
literacy are tied to early language development in children who have not yet learned
to read.

White-Schwoch et al. (2015) conducted a second study of phonological pro-
cessing in preschoolers. Unlike previous experiments that focused on one aspect
of the FFR, they combined three facets of neurophysiological processing, each of
which is tied to literacy in older children (see Sect. 6.4.2.2): timing, representa-
tion of formant features, and consistency. They developed a statistical model
incorporating these features in response to a CV transition in noise and found that
the model was strongly predictive of children's phonological processing.
Additionally, they could forecast how children perform on an array of early
literacy tests one year later, providing the first longitudinal evidence that the FFR
predicts the development of language skills. Finally, they applied this model to
school-aged children and found it reliably identifies which children have a
learning disability, supporting the idea that the FFR could be a clinical tool to
identify children at risk for learning disabilities (see Anderson, Chap. 11 for a
discussion of FFR clinical applications).

6.4.3 Interim Summary

Together, the studies discussed in the previous sections show how FFRs are distinct
in children with language-based LPs, how FFRs elucidate mechanisms underlying

communication impairments in those populations, and how FFRs predict future success and struggle in everyday communication skills. A common thread through these studies is that listening and language are connected. Specifically, the ability to understand speech in noise is partially connected to language skills, and FFR signatures of those domains partially overlap (Fig. 6.4). Hearing in noise pulls distinctly on the ability to group talkers and auditory objects and is characterized by FFR markers of those skills. Language and literacy skills pull distinctly on the ability to categorize phonemes and are characterized by FFR markers of those skills. Both rely on a resiliency to difficult listening situations, accurate distinctions of sounds, and sensitivity to stimulus context. This work sets the stage for studies of long-term auditory experience, because these same neurophysiological processes are affected by that experience.
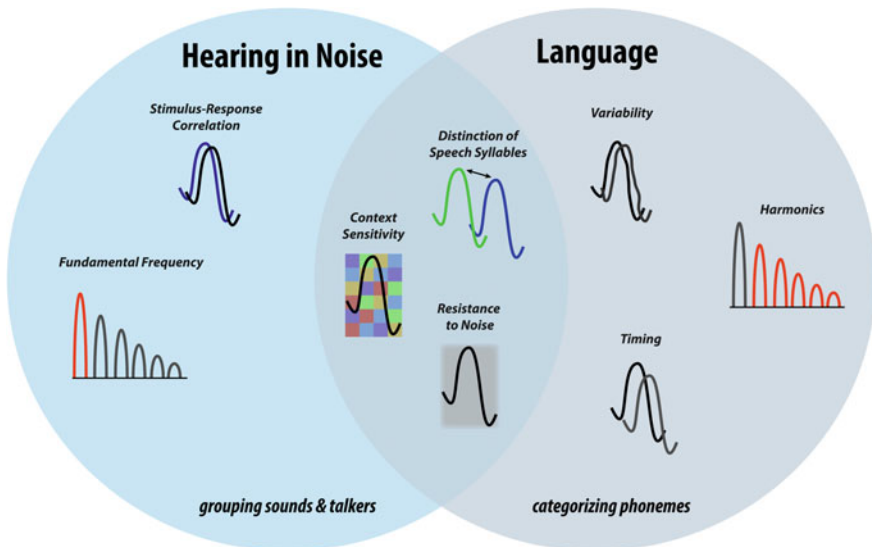

Fig. 6.4 Neural signatures of everyday communication: the listening-language overlap. Language skills are tied to listening, particularly the ability to understand speech in noise, but these neural signatures only partially overlap. The ability to understand speech in noise requires grouping sounds and talkers, facilitated by strong representation of the fundamental frequency and processing of meaningful stimulus ingredients. Language development (particularly language skills important for literacy) requires accurate categorization of incoming phonemes, facilitated by strong harmonic representation, fast timing, and a healthy base of consistent auditory processing. Both domains are supported by a sensitivity to stimulus context, accurate distinctions of contrastive speech sounds, and resiliency to the degrading effects of background noise on speech processing. (Original figure by White-Schwoch and Kraus)

6.5   FFR Studies of Music Training

The literature on music training vis-à-vis the FFR is large, and many converging experiments reinforce the evidence discussed below. Strait and Kraus (2014) provide a comprehensive review on the topic, couched in a broader literature on music and brain plasticity.

6.5.1   Music Training as a Model for Auditory Learning

Music engages a rich and diverse series of brain networks. The literature on music training and brain plasticity is vast and diffuse with respect to sensory modalities, cognitive functions, and methodology. Although the emphasis here is on FFR studies of music training, these should be considered in the broader context of changes in auditory anatomy and physiology attributed to music training, ranging from the cochlea to cortex. In these experiments, "musicians" are individuals who regularly (at least 2–3 times/week) produce music by instrument or voice. In general, they exclude individuals who engage solely in music listening or music appreciation, both of which appear insufficient to spark neural remodeling (Kraus et al. 2014a). And while many studies are correlational, random-assignment experiments reinforce a causal role for music training in conferring this plasticity (Chobert et al. 2012; Kraus et al. 2014b). Importantly, these neurophysiological enhancements are not merely parlor tricks: as summarized in Fig. 6.5, musicians enjoy a host of advantages for auditory cognition and perception over their peers.

The cognitive-sensorimotor-reward model of auditory learning (Sect. 6.3.2) is encapsulated by music training. The *sensorimotor* component of music making is perhaps the most self-evident, but what is thought to be important in the context of music and learning is the integration across sensory modalities (Limb and Braun 2008). Imagine that a piano player must continuously listen to sounds that are produced, use auditory feedback to modulate motor control, and use visual information, such as cues from other musicians. The *cognitive* component includes directed attention that frequently shifts between auditory streams (for example, rapidly switching listening from violins to trombones) and strong engagement of auditory working memory systems. The *reward* component comes into play given the extraordinary activation of limbic systems during music performance and listening (Blood et al. 1999). Classic studies on auditory cortex plasticity show that pairing limbic system stimulation with auditory cues catalyzes neuroplasticity (Bakin and Weinberger 1996; Kilgard and Merzenich 1998). Thus, the auditory-reward coupling inherent in music making, combined with top-down cognitive modulation, creates ideal circumstances for auditory learning (Patel 2011).
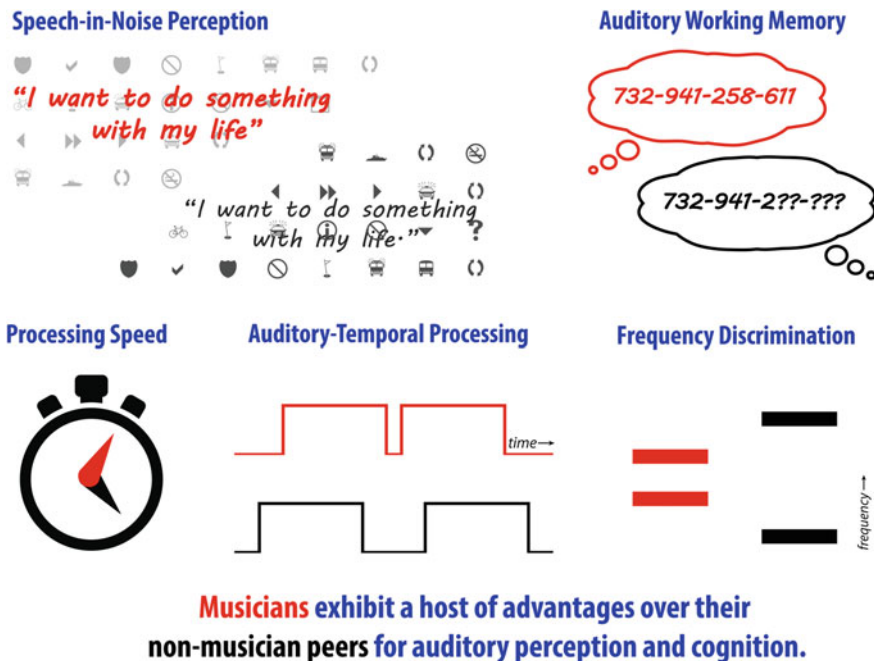
Speech-in-Noise Perception

"I want to do something with my life"

"I want to do something with my life."

Auditory Working Memory

732-941-258-611

732-941-2??-???

Processing Speed

Auditory-Temporal Processing

time→

Frequency Discrimination

frequency →

Musicians exhibit a host of advantages over their non-musician peers for auditory perception and cognition.

Fig. 6.5 Music training strengthens auditory perception and cognition. Evidence across their life spans shows that musicians outperform their nonmusician peers on tests of auditory function. These include (*clockwise from top left*) speech perception in noise, auditory working memory, frequency discrimination, auditory-temporal processing, and processing speed. (Original figure by White-Schwoch and Kraus)

6.5.2 *FFRs are Distinct in Listeners with Music Experience*

Musacchia et al. (2007) conducted the first study of music training and the FFR. They measured FFRs to a CV syllable and to a cello sound in audio-only and audiovisual conditions. Across stimuli and listening conditions, musicians had faster and larger responses than their peers. Additionally, FFR amplitudes were correlated to the extent of music training in both audio-only and audiovisual conditions, reinforcing the idea that neurophysiological enhancements observed in musicians are products of experience, not preexisting differences.

The second landmark study of music training and the FFR was inspired by Krishan and colleagues' seminal work showing that experience with a tone language enhances FFR tracking of pitch contours (Krishnan and Gandour, Chap. 3). Wong et al. (2007) asked if music training affects this linguistic processing as well. They found that amateur musicians had more accurate FFR tracking of pitch contours than nonmusicians. Additionally, the extent of music training correlated to the accuracy of pitch tracking, suggesting that additional music training heightens

neurophysiological acuity for linguistically relevant sounds, which is consistent with the idea that neural circuits for music and language overlap (Patel 2008).

Bidelman et al. (2011) replicated and extended this work, comparing musicians to tone language speakers (native Mandarin speakers). Although musicians and Mandarin speakers both had stronger tracking of tone contours and iterated rippled noise than English-speaking nonmusicians, musicians specialized for more musical sounds and Mandarin speakers specialized for more linguistic sounds. Thus, although there is a cross-domain transfer between music and language, specialization still occurs along dimensions that are the most behaviorally salient for a listener (see Sect. 6.3.4). Strait et al. (2012a) tested a similar concept by comparing two groups of musicians—pianists and non-pianists—all with a similar degree of training. Pianists had stronger FFR coding of the timbre of a piano than of bassoon or tuba sounds, whereas non-pianists had equivalent coding across instruments. This supports the notion of "specialization among the specialized" in experience-dependent plasticity: auditory expertise is a continuum that is amplified by further training.

Strait et al. (2009) pursued both the questions of selective modulation of auditory function and of the generalization of music training to other kinds of sounds by eliciting FFRs to a baby cry (an emotional sound) in musicians and nonmusicians. The stimulus was important because it contained spectrally simple (periodic) and complex (containing acoustic transients) segments. Musicians had smaller responses to the simple segment, but larger responses to the complex segment. This was interpreted to reflect both attenuation of neural processing for simple sounds (a neural economy of resources) and enhancement of neural processing for challenging sounds (neural recruitment to process difficult stimuli). This study reinforces that experience-dependent plasticity is not an "all-or-none" phenomenon (see Sect. 6.3.4).

It should be noted that the FFR literature on music training reinforces the view of the auditory system as a distributed, but integrated, circuit (see Sect. 6.2). It has been argued that music training influences subcortical processing through the corticofugal system and that this integration heightens neural acuity throughout the system (Kraus and Chandrasekaran 2010). This is evident in studies that consider the interactions between different levels of auditory processing. For example, Lee et al. (2009) show that musically relevant cochlear nonlinearities are represented more robustly in the midbrain FFRs of musicians. Additionally, music training strengthens the coherence between FFRs and cortical responses in younger (Musacchia et al. 2008) and older adults (Bidelman and Alain 2015).

Do FFR signatures of music training generalize to everyday communication? With respect to auditory cognition and perception, there is evidence that musicians have superior speech perception in noise, processing speed, auditory working memory, auditory-temporal processing (such as backward masking thresholds), and frequency discrimination (Fig. 6.5). Longitudinal and cross-sectional studies argue that these cognitive and communicative gains are intimately tied to augmented

auditory processing (Strait and Kraus 2014). Literacy skills and the ability to listen to speech in noise are emphases in the music-communication overlap (Kraus and White-Schwoch 2016).

Are these enhancements linked to the FFR? Weiss and Bidelman (2015) addressed this question in the context of speech perception in an elegant experiment. They measured FFRs to sounds along a continuum from the vowels /u/ to /a/. They "sonified" the FFRs—transformed the waveforms into audible sound files, capitalizing on the spectotemporal richness of FFRs—then used the sonified FFRs as stimuli in a simple categorical perception task. Listeners categorized the FFRs of musicians more quickly and more accurately than the FFRs of nonmusicians. This suggests that the stronger auditory processing by musicians, particularly of speech in noise (Parbery-Clark et al. 2009b; Zendel and Alain 2012), may be grounded, in part, in processes reflected by the FFR.

6.5.3 Music Training Affects Speech Processing Across the Life Span

6.5.3.1 Preschoolers

Strait et al. (2013) investigated the auditory-neurophysiological impact of preschool music classes by eliciting FFRs to a consonant-vowel syllable in quiet and noise. Preschoolers (ages 3–5 years) engaged in music classes had faster responses to the onset and consonant transition of speech in quiet and noise, but the groups had identical response timing for a vowel. Additionally, these children had stronger stimulus-response correlations, indicating more accurate neural coding of speech, and their responses were more resilient to background noise. Moreover, preschoolers had stronger neurophysiological distinctions of contrastive syllables (Strait et al. 2014); because these FFR metrics are linked to literacy (see Sect. 6.4.2.3), including in preschoolers, these results support the idea that music training boosts communication skills. In contrast to older children and adults engaged in music lessons, preschoolers did not exhibit enhanced processing of speech harmonics (see Sects. 6.5.3.2 and 6.5.3.3).

A subset of children returned after one additional year of music training, providing the first longitudinal study of music training vis-à-vis the FFR. Preschoolers exhibited even stronger neurophysiological resiliency to the degrading effects of background noise on response timing after the additional year of training. These results corroborate a causal role for music training in engendering brain plasticity. In the context of studies on music training across the life span, these results suggest that the musician's neural signature emerges progressively (Fig. 6.6).
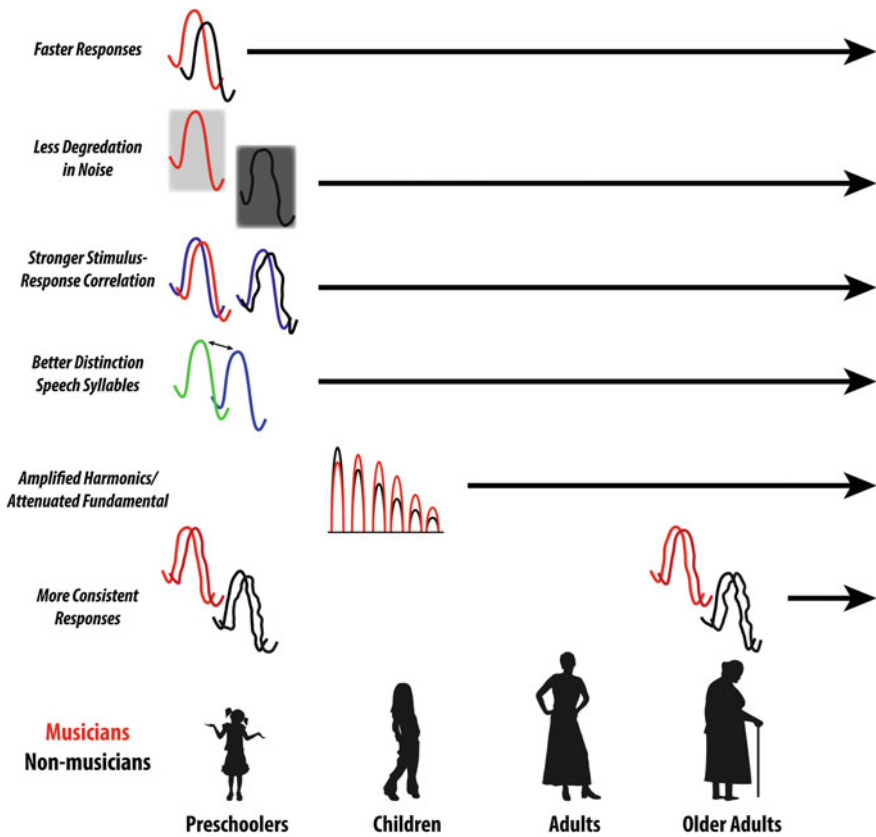
Fig. 6.6 Musicians exhibit a neural signature of their experience that is manifest in the FFR and emerges progressively across the life span. Preschool musicians have faster responses, more resiliency to background noise, stronger stimulus-response correlations, and better neurophysiological distinctions of contrastive speech syllables (see Fig. 6.3). Older children begin to show stronger representation of speech harmonics. The last element of the musician's neural signature emergences in senescence, when musicians have more consistent responses than their nonmusician peers. (Original figure by White-Schwoch and Kraus)

6.5.3.2   Children

Strait et al. (2012b) conducted a complementary series of studies in older children (ages 7–13 years), employing identical techniques as in preschoolers (see Sect. 6.5.3.1). Children with music training had faster responses to the onset and CV transition of speech presented in quiet and noise; they also exhibited more resiliency to background noise and stronger stimulus-response correlations. Similar to preschoolers, these older children with music training exhibited stronger neurophysiological distinctions of contrastive speech syllables (Strait et al. 2014). Additionally, children with music training had stronger processing of speech

harmonics, suggesting this aspect of the musician's neural signature emerges following additional music training.

Strait and colleagues also investigated auditory behaviors, including sentence-in-noise perception, auditory working memory, and auditory attention. Consistent with evidence from adults (Fig. 6.5), they found that children with music training outperformed their peers on these tests. Additionally, a child's performance on these tests related to the integrity of neural processing (including response timing and harmonic representation). This again corroborates the idea that strengthening neurophysiological processing boosts skills important for everyday communication.

6.5.3.3 Young Adults

Parbery-Clark et al. (2009a) conducted a study of young adults (ages 19–30 years) and measured FFRs to speech in quiet and noise. Young adult musicians had faster responses to the onset and CV transition, stronger stimulus-response correlations, more resiliency to background noise, stronger representation of speech harmonics, and stronger neurophysiological distinctions of speech syllables. All of these facets of neurophysiological processing related to the ability to understand sentences in noise, again corroborating the idea that music training engenders neurophysiological enhancements to speech processing that generalize to everyday communication skills.

In a later study, Parbery-Clark et al. (2013) evaluated the extent to which young adults with music training take advantage of binaural hearing, which has a recognized role for listening in noise. They measured FFRs to /d/ in quiet monaurally to each ear and diotically (same sound to both ears). They chose a sound for which a musician advantage had not been observed and found that while the groups had similar responses to monaural input, musicians had faster and less variable responses to the diotic input (Fig. 6.7). Moreover, musicians had a larger binaural "benefit" in terms of the acceleration of FFR timing and decrease of response variability, suggesting a musician advantage may be rooted in binaural processing.

6.5.3.4 Middle-Aged and Older Adults

Parbery-Clark et al. (2012a, b) adopted similar techniques to study auditory processing in middle-aged and older adults with music training (ages 45–65 years). A hallmark of auditory aging is loss of temporal precision that disrupts the ability to understand rapidly changing speech sounds and the ability to understand speech in noisy environments (for a review see Gordon-Salant 2014). In the FFR, aging manifests as a diverse and wide-ranging neural signature, including a prolongation of response timing in response to consonant transitions (see Anderson, Chap. 11). Parbery-Clark et al. (2012a) showed that older adult musicians have identical
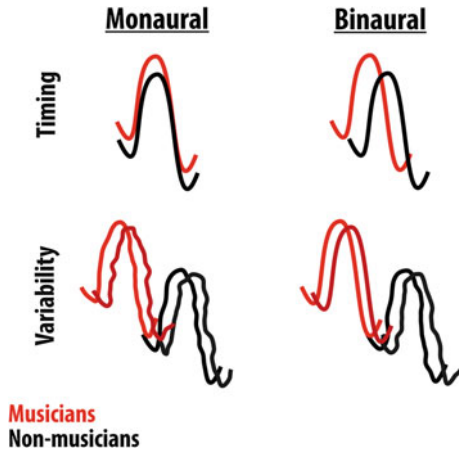
Fig. 6.7 Young adult musicians have stronger binaural responses to speech. Parbery-Clark et al. (2013) compared monaural and binaural FFRs in young adult musicians and nonmusicians. The two groups had similar *monaural* responses with respect to timing and consistency. However, musicians had faster and more consistent *binaural* responses tied to superior speech perception in noise, suggesting music experience may strengthen binaural processing. (Original figure by White-Schwoch and Kraus)
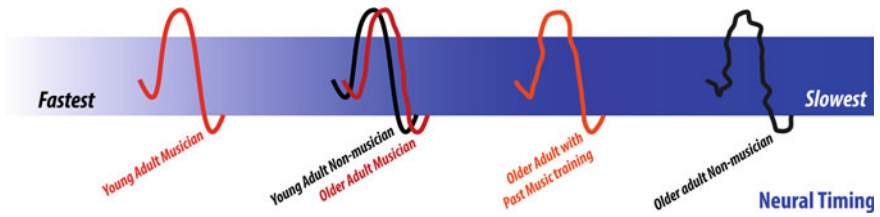

Fig. 6.8 Music training affects neural timing. FFRs to consonant-vowel transitions are faster in musicians than in nonmusicians. The fastest responses are observed in young adult musicians. Older adults with a life of music training have identical response timing as young adult nonmusicians. Older adults with music training early in life (but not for 50+ years) still have faster responses than older adult nonmusicians. (Original figure by White-Schwoch and Kraus)

response timing to young adult nonmusicians, suggesting that lifelong music training offsets this signature of aging (see Fig. 6.8). Bidelman and Alain (2015) also reported that older adult musicians exhibited fewer age-related declines in neural processing, and that older musicians had tighter integration between cortical and subcortical processing than their peers.

In a subsequent study, Parbery-Clark et al. (2012b) compared older adults with music training to older adult nonmusicians, and found that the musician's signature persists in this older population. Older adults with continuous, lifelong music training had faster responses, more robust responses to speech harmonics, stronger stimulus-response correlations, and more resiliency to background noise. Again,

neurophysiological benefits were concomitant to gains in speech perception in noise and auditory working memory. Interestingly, a new aspect of the musician signature arose: older adult musicians had less variable neurophysiological responses, off-setting an additional element of the FFR signature for aging (Anderson et al. 2012).

6.5.4 Music Training Imparts an Enduring Biological Legacy

Does music training have to be lifelong to engender biological enhancements to sound processing? The simple answer is no, although the neurophysiological profile in listeners with past music experience is more circumscribed.

Skoe and Kraus (2012) tackled this question in young adults with a history of music training who did not currently play an instrument. They measured FFRs to a series of triangle tones and found that young adults with the most prior music training had the largest response to the F_0 of the tones, followed by a group with a small amount of music training, followed by a group with no music training. Thus, it appears that past auditory experiences, even if they are not maintained, still influence auditory function.

White-Schwoch et al. (2013) addressed a similar question in older adults. They also compared three groups (one with a modest amount of training, one with a little, and one with none). However, they considered individuals who had not played music for 50+ years. Nevertheless, White-Schwoch and colleagues found a systematic relationship between the extent of music training and the speed of FFRs to a consonant transition: older adults with the most music training had the fastest responses (Fig. 6.8). This is noteworthy because it ties into one of the hallmarks of aging (see Sect. 6.5.3.4 and Anderson Chap. 11). However, no other aspects of the aging musician signature were apparent, suggesting that while music training during childhood imparts a lifelong biological legacy, that legacy is not as pronounced as in lifelong musicians.

Three hypotheses emerge to explain these findings. The first hypothesis is that innate differences in auditory abilities drive some individuals to pursue music for longer than others. For example, an individual with faster FFRs may experience more success playing an instrument. It is unlikely, however, that there is a direct relationship between FFR timing and the number of years an individual feels driven to make music, suggesting a causal role for music training in engendering these changes. Additionally, longitudinal studies show a causal role for music training in shaping auditory physiology (Chobert et al. 2012; Kraus et al. 2014b).

The second hypothesis is that early experience tunes the auditory system during childhood and these changes are fixed. This possibility is also unlikely, however, given the potential for learning to shape auditory functions throughout life (see Carcagno and Plack, Chap. 4 for a discussion of short-term auditory training in

adults). Thus, a third hypothesis emerges: music training teaches a listener to interact more meaningfully with sound, making everyday auditory activities learning experiences. By directing attention to sound in cognitively demanding and rewarding contexts (see Sect. 6.3.2), music training may facilitate sound-to-meaning connections during daily listening, an idea evocative of meta-plasticity: neuroplasticity that is a consequence of neuroplasticity (Abraham 2008).

6.6 FFR Studies of Socioeconomic Status

The literature on SES and the FFR is in its infancy. Nevertheless, early evidence supports the contention that growing up in poverty negatively affects auditory brain function in ways that are both similar and distinct to neural profiles of populations with communication impairments (see Sect. 6.4).

6.6.1 Socioeconomic Status as a Model for Auditory Learning

The idea that SES provides a model for auditory learning is rooted in the concept that a child's everyday sensory milieu shapes brain structure and function (see Sect. 6.3.1). The maturation-by-experience hypothesis predicts divergent outcomes for children raised in relatively high versus low socioeconomic environments. Much like music, the socioeconomic environment includes a large number of factors that are candidates for driving this plasticity and, much like music, it is thought that SES engenders plasticity through this confluence. Socioeconomic status is a broad term reflecting factors that are difficult to quantify. Chief among these is an impoverished linguistic repertoire that may impede children's mapping of sounds in the environment to meaning. Many studies (including those using the FFR) use maternal education as a proxy for SES (Hoff et al. 2012).

Hart and Risley (1995) conducted a landmark study of toddlers' home language environments. Children in low SES households heard 30 million fewer words by three years of age than their peers. Additionally, these children heard 40% the number of different words as their peers. Thus, these children grow up in homes that are linguistically impoverished with respect to both the quantity and quality of their daily exposure. The latter is often forgotten in popular accounts of this seminal finding but is likely to be a major factor. After all, a mother's voice is one of the most rewarding stimuli available to a child (Abrams et al. 2016) and may facilitate learning a mother's linguistic repertoire even if said repertoire is impoverished. This scenario reinforces an important idea of neuroplasticity: the conditions may be correct to facilitate learning, but what is learned may be maladaptive (Kilgard 2012).

The physical environments of low SES neighborhoods provide another cluster of factors that may shape auditory function. Low SES neighborhoods run the risk of higher levels of toxins such as poor water quality and air pollution. Additionally, low SES neighborhoods may be noisier (imagine housing near an airport); background noise interferes with auditory cortical map development (Chang and Merzenich 2003) and degenerates afferent synapses at the inner hair cells (Sergeyenko et al. 2013; also see Shinn-Cunningham, Varghese, Wang, and Bharadwaj, Chap. 9 for a discussion of environmental noise exposure and the FFR). Finally, poor nutrition may impede brain development. Although there are no direct relationships established between any of these factors and the FFR, the integrative model of the auditory system supports the hypothesis that this confluence is reflected by the FFR.

Together, the constellation of factors associated with SES cover cognitive domains (such as the sophistication of language input and educational opportunities) and sensorimotor domains (such as environmental noise exposure and the total amount of auditory input) in rewarding settings (such as hearing mothers' voices). As in the previous discussions of the role of music, the emphasis of this chapter is studies of SES vis-à-vis the FFR. Note this only provides a glimpse into a large literature concerned with the development of brain structure and function in low SES populations (Stevens et al. 2009; Noble et al. 2015) and educational and socioemotional outcomes (Leventhal and Brooks-Gunn 2000).

6.6.2   The Neural Signature of SES

Skoe et al. (2013) established the neural signature of SES (Fig. 6.9) by studying adolescents, whom they dichotomized into groups with lower levels of maternal education (high school or less) and higher levels of maternal education (any amount of postsecondary schooling). Children in the low maternal education group had less stable FFRs and poorer representations of speech harmonics. Additionally, this group had "noisier" neural activity; that is, amplitudes were higher during the brief intervals of silence between each presentation of the [da] in the recording session. This increase in spontaneous neural activity may be likened to static on the telephone, and Skoe et al. (2013) suggested it may interfere with precise processing. Moreover, the low maternal education group performed more poorly than their peers on tests of reading and auditory working memory.

A noteworthy aspect of this study is that differences in academic achievement and neurophysiological processing were evident despite a great deal of homogeneity between the groups, which were matched for age, sex, intelligence, hearing, and even came from the same schools. Although the groups differed in the extent of maternal education, this difference was small (≈3.5 years). Nevertheless, there were clear group differences. This suggests that modest disparities in maternal education, and by extension SES, cascade to tangible differences in auditory neurophysiology and skills important for everyday communication.
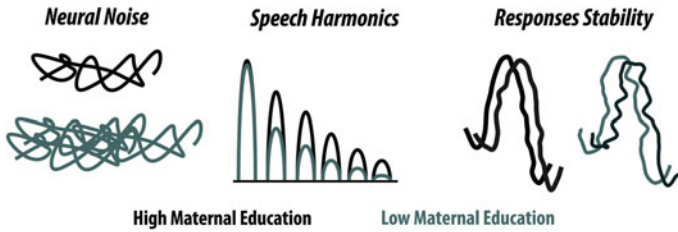
Fig. 6.9 The neural signature of socioeconomic status (SES). Children whose mothers have low levels of maternal education (a proxy for low SES) have distinct neural processing from their peers. Their FFRs have increased levels of neural noise, exhibit poor representation of speech harmonics, and are less stable across trials. (Original figure by White-Schwoch and Kraus)

6.7 Music Meets SES

There is broad interest in the use of community music programs to enrich perceptual, cognitive, and emotional development in underserved populations, including low SES children. Two groups of studies have used the FFR as an outcome measure to determine if music sparks neuroplasticity in low SES populations.

6.7.1 Music Training Catalyzes Neuroplasticity in Low SES Populations

6.7.1.1 Children

The first group of studies involved the Harmony Project (www.harmony-project.org), a mentorship organization in Los Angeles that provides free music instruction to children from gang-reduction zones. Children (ages 6–9 years at study onset) were randomly assigned to engage in instrumental music lessons or to spend a year on a waitlist, with guaranteed admission to music lessons one year later. Kraus et al. (2014b) found that the neurophysiological distinction of contrastive syllables (see Fig. 6.3) was stronger following two years of music training, but one year was insufficient. In a follow-up study, they pitted instrumental music training against music appreciation classes and found that children who actively made music had faster responses and stronger representation of speech harmonics (Kraus et al. 2014a).

This cluster of neurophysiological measures is evocative of the neural signature of literacy and language disorders (see Sect. 6.4.2.2) and provides additional evidence that music training may support the development of these skills or, at a minimum, their neural substrates (for reviews see Tierney and Kraus 2014;

Kraus and White-Schwoch 2016). This work extends this idea to low SES populations, which are known to exhibit lower levels of academic achievement (see Sect. 6.6.2). In fact, children engaged in instrumental music instruction performed better on tests of rhythmic aptitude (Slater et al. 2013) and rapid automatized naming (Slater et al. 2014)—two skills thought to support reading. Additionally, these children had better perception of sentences in noise (Slater et al. 2015), providing longitudinal evidence supporting the aforementioned correlational findings (see Sect. 6.5.1).

These experiments illustrate two principles of auditory learning. First, it takes time to change the brain. One year of music training was insufficient to spark neural remodeling, suggesting that the auditory system has some resistance to changing its automatic response properties. This is not to say that auditory processing cannot adapt to new contexts (see Escera, Chap. 5); rather, the deeply ingrained response properties evident during passive listening situations reflect the consequences of prolonged and repeated auditory experience (see Sect. 6.3.3). Second is that *making music matters*, that is, active engagement during the learning phase is necessary to engender remodeling evident in the learned phase.

6.7.1.2 Adolescents

The second group of experiments was conducted in collaboration with the Chicago Public Schools and involved a group of adolescents (starting age ≈ 14 years). The study focused on children in charter schools that offered co-curricular enrichment programs, either incorporating music training (instrumental lessons or choir) or paramilitary training (Junior Reserve Officer's Training Corps, JROTC, emphasizing discipline and athletics). Although these experiments did not employ random assignment (children chose their training), including the JROTC group provides an active control matched for the amount of time spent engaged in enriching activates. Thus, this work may arbitrate between potential effects of music training sensu stricto and childhood enrichment more generally.

Once again a single year of music training was insufficient to prompt neurophysiological changes, but after two years adolescents engaged in music training had faster responses to speech in noise (Tierney et al. 2013), consistent with the musician's neural signature (see Sect. 6.5.2). After a third year of training, the music group showed less variable FFRs than their peers (Tierney et al. 2015). Additionally, cortical responses to speech were more mature in children with music training. These results harken back to the maturation-by-experience hypothesis by showing that auditory enrichment piggybacks on neurophysiological processes still undergoing a phase of late maturation. In addition, the results are consistent with the view that altering a sensory milieu cascades to changes in sound processing. Indeed, adolescents engaged in music training also showed stronger phonological processing than their peers, a key literacy skill that is linked to FFR variability (White-Schwoch et al. 2015).

6.7.2   SES Shapes Auditory Processing Directly and Indirectly

Anderson et al. (2013) considered SES in a study of speech perception in noise. They studied 120 older adults and used statistical modeling to determine the influences of auditory-neurophysiological processing (FFRs to [da] in noise), cognition (auditory working memory and attention), hearing (air-conduction thresholds and distortion product otoacoustic emissions), and life experience (SES, past music training, and physical activity) on the ability to understand speech in noise. Life experience (including SES, measured by maternal education and the subjects' own educational attainment) influenced the ability to understand speech in noise with an additional indirect influence via neurophysiological processing. In other words, life experiences shaped the integrity of the FFR, which in turn shaped the ability to understand speech in noise. Moreover, like Skoe et al. (2013), they found a direct link between SES and formant processing (see Sect. 6.6.2).

Additionally, Anderson et al. (2013) considered whether their subjects underwent music training at any point during their lives. Older adults with music experience—regardless of whether or not they still played—relied more heavily on cognitive factors to understand speech in noise, consistent with the idea that music training shapes a listener's everyday relationship with sound (see Sect. 6.5.4). However, the influence of SES belied the apparent effects of music training (Fig. 6.10). This is not surprising, as families with more resources may be more likely to engage in music training.

The compounding influence of SES and music training highlights how SES is a complex and multifactorial phenomenon with both direct and indirect effects of auditory neurophysiology. For example, noise toxicity in impoverished neighborhoods may directly affect the auditory system, but the lack of opportunities for auditory enrichment, which is itself a consequence of low SES, also likely affects auditory processing.

6.7.3   SES and Bilingualism: A Complementary Case of Auditory Enrichment

The majority of studies discussed above asked how one sort of life experience, such as music or low SES, affects auditory function. Krizman et al. (2016) asked how experience with sound—bilingualism—interacts with SES. They tested a population of Spanish-English bilingual adolescents and compared them to a matched group of English-monolingual adolescents, building on work establishing that Spanish-English bilinguals have enhanced FFRs to the F_0 and more consistent FFRs to vowels (Krizman et al. 2012). Next, they split the two groups according to SES, using a similar strategy as Skoe and colleagues (see Sect. 6.6.2).
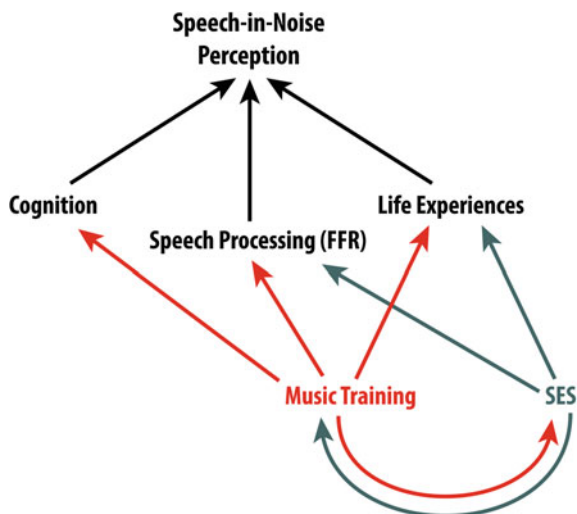
Fig. 6.10 Socioeconomic status (SES) influences auditory functions directly and indirectly. The ability of older adults to understand speech in noise is shaped by cognitive factors, speech processing (FFR), and life experience (music training and SES). Music training enhances cognition and speech processing, and SES affects speech processing. However, the apparent influence of music training could be attributed to SES. Individuals from lower-SES backgrounds were less likely to have received music training at all, pointing to the complex interactions between multiple types of auditory experience in predicting communication skills. Thus, music and SES interact to shape auditory processing and everyday communication. (Adapted from Anderson et al. 2013)

Bilingualism buffered the effects of SES on FFR stability. Specifically, the responses of low SES bilinguals were as consistent as high-SES monolinguals. Thus, enriching an auditory environment (by learning a second language) may counteract some of the consequences of growing up in impoverished environments. Presumably this buffer specifically addresses the linguistic impoverishment endemic to low SES children (see Sect. 6.6.1) and shows how enrichment and deprivation interact to shape auditory function. A similar finding comes from Zhu et al. (2014), who raised a group of rats in environmental deprivation but found that intense perceptual training as adults restored auditory cortical processing.

This also illustrates that the selective modulation(s) conferred through auditory learning may interact with modulations following other experiences. Practically, this suggests that therapeutic interventions might reverse some of the neural signatures of deprivation and communication disorders (see Carcagno and Plack, Chap. 4). Together, the bilingualism and musician models of auditory learning provide complementary evidence that auditory enrichment counteracts auditory deprivation.

6.8   Summary: Time Traveling Through the Auditory System

The auditory system is inherently plastic but strikes a remarkable balance between a propensity for, and a resiliency against, change. Everyday experiences in and through sound push the auditory system to reorganize, and the FFR provides an approach to explore this reorganization in humans. This reorganization bears on communication skills. An individual's FFR provides a window into his or her communication abilities, and signature patterns of FFR properties characterize several groups of listeners. These include individuals with language problems (such as dyslexia, SLI, APD, and ASD), and listeners who have undergone extensive experiences, such as musicians or children from low SES backgrounds. The latter have been the emphasis of this chapter and provide two experimental models of auditory learning in humans.

An intriguing notion emerges when considering this work as a whole: like Janus, an individual's FFR is a window to the past and doorway to the future, and thus offers a form of time travel through the auditory system. Each person's life in sound is a push and pull between enhancements and diminutions to sound processing, meaning FFR properties indicate the imprint of past experiences, good or bad (Fig. 6.11). The same response properties predict what the future holds for that individual's ability to communicate. The FFR provides this remarkable window to the past, present, and future because it is a measure of integrative auditory function. But diverse research on the FFR shows that the auditory future is not set in stone: auditory experiences that boost a relationship to sound through cognitive,
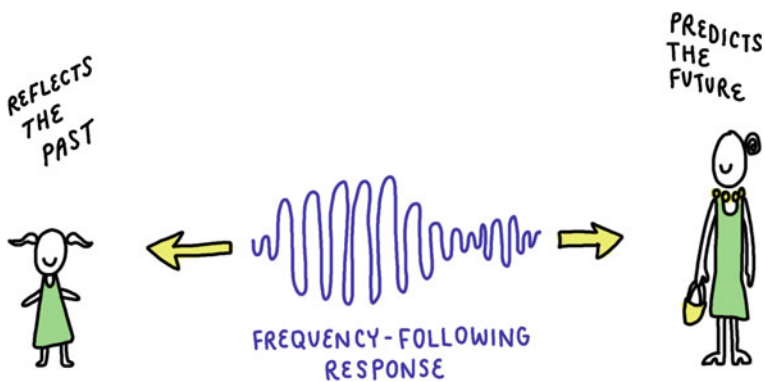

Fig. 6.11 Time traveling through the auditory system. Auditory anatomy and physiology reflect the legacy of a life in sound and predict the consequences of these experiences for future communication skills. Thus, the FFR can be thought of as a vehicle for time travel. An individual's FFR, measured today, is a biological mosaic that reflects the past and can objectively predict future success and challenges in everyday communication such as language and listening in noise. (Image by Katie Shelly)

sensorimotor, and reward networks can ameliorate poor communication skills and set up a new path for auditory processing.

This idea represents the next step for FFR research in learning and communication. How do multiple sorts of experiences interact to shape communication? Can adaptive learning completely offset maladaptive learning, or does it at least compensate for reduced function (e.g., Chambers et al. 2016)? Given promising evidence that the FFR can predict future communication impairment, can it also be used to motivate targeted early interventions? These questions can be explored as FFR techniques advance. As new neural signatures of communication crystallize (e.g., concussion; Kraus et al. 2016), these signatures can be used to fit individuals to different patterns. For example, one child may show a pattern that motivates music training and another may show a pattern that motivates second language training. As the FFR becomes more widely available, and its underlying mechanisms more thoroughly understood, it has exciting potential as a clinical tool to aid human communication. Moreover, it has the potential to bridge auditory neuroscience with education and social policy, both by highlighting the importance of everyday auditory experiences in development and well-being and by providing strategies in their support.

Compliance with Ethics Requirements Travis White-Schwoch declared that he had no conflict of interest.

Nina Kraus is Chief Scientific Officer of Synaural, a company working to develop a user-friendly measure of auditory processing.

References

Abraham, W. C. (2008). Metaplasticity: Tuning synapses and networks for plasticity. *Nature Reviews Neuroscience, 9*(5), 387.

Abrams, D. A., Nicol, T., Zecker, S. G., & Kraus, N. (2006). Auditory brainstem timing predicts cerebral asymmetry for speech. *The Journal of Neuroscience, 26*(43), 11131–11137.

Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2011). A possible role for a paralemniscal auditory pathway in the coding of slow temporal information. *Hearing Research, 272*(1), 125–134.

Abrams, D. A., Chen, T., Odriozola, P., Cheng, K. M., et al. (2016). Neural circuits underlying mother's voice perception predict social communication abilities in children. *Proceedings of the National Academy of Sciences of the U S A, 113*(22), 6295–6300.

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences, 11*(11), 458–465.

Ahissar, E., Sosnik, R., & Haidarliu, S. (2000). Transformation from temporal to rate coding in a somatosensory thalamocortical pathway. *Nature, 406*(6793), 302–306.

Anderson, S., & Kraus, N. (2016). Auditory processing disorder: Biological basis and treatment efficacy. In C. G. Le Prell, E. Lobarinas, A. N. Popper, & R. R. Fay (Eds.), *Translational research in audiology, neuro-otology, and the hearing sciences*. Springer Science + Business Media: Berlin.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2012). Aging affects neural precision of speech encoding. *The Journal of Neuroscience, 32*(41), 14156–14164.

Anderson, S., Skoe, E., Chandrasekaran, B., & Kraus, N. (2010a). Neural timing is linked to speech perception in noise. *The Journal of Neuroscience, 30*(14), 4922–4926.

Anderson, S., Skoe, E., Chandrasekaran, B., Zecker, S., & Kraus, N. (2010b). Brainstem correlates of speech-in-noise perception in children. *Hearing Research, 270*(1–2), 151–157.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hearing Research, 300*, 18–32.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2015). Development of subcortical speech representation in human infants. *The Journal of the Acoustical Society of America, 137*(6), 3346–3355.

Atiani, S., David, S. V., Elgueda, D., Locastro, M., et al. (2014). Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron, 82*(2), 486–499.

Ayala, Y. A., Udeh, A., Dutta, K., Bishop, D., et al. (2015). Differences in the strength of cortical and brainstem inputs to SSA and non-SSA neurons in the inferior colliculus. *Scientific Reports, 5*(10383), 2016. Doi:10.1038/srep10383

Bajo, V. M., Nodal, F. R., Moore, D. R., & King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature Neuroscience, 13*(2), 253–260.

Bakin, J. S., & Weinberger, N. M. (1996). Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. *Proceedings of the National Academy of Sciences of the U S A, 93*(20), 11219–11224.

Banai, K., Hornickel, J., Skoe, E., Nicol, T., et al. (2009). Reading and subcortical auditory function. *Cerebral Cortex, 19*, 2699–2707.

Basu, M., Krishnan, A., & Weber-Fox, C. (2010). Brainstem correlates of temporal auditory processing in children with specific language impairment. *Developmental Science, 13*(1), 77–91.

Behrmann, M., & Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual recognition. *Trends in Cognitive Sciences, 17*(5), 210–219.

Benasich, A. A., & Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioural Brain Research, 136*(1), 31–49.

Bidelman, G. M., & Alain, C. (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *The Journal of Neuroscience, 35*(3), 1240–1249.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23*(2), 425–434.

Bishop, D. V. (1997). *Uncommon understanding (classic edition): Development and disorders of language comprehension in children*. New York: Psychology Press.

Bishop, D. V., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry, 31*(7), 1027–1050.

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience, 2*(4), 382–387.

Centanni, T. M., Booker, A., Sloan, A., Chen, F., et al. (2014). Knockdown of the dyslexia-associated gene *Kiaa0319* impairs temporal responses to speech stimuli in rat primary auditory cortex. *Cerebral Cortex, 24*(7), 1753–1766.

Chambers, A. R., Resnik, J., Yuan, Y., Whitton, J. P., et al. (2016). Central gain restores auditory processing following near-complete cochlear denervation. *Neuron, 89*(4), 867–879.

Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., & Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: Implications for developmental dyslexia. *Neuron, 64*(3), 311–319.

Chang, E. F., & Merzenich, M. M. (2003). Environmental noise retards auditory cortical development. *Science, 300*(5618), 498–502.

Chobert, J., François, C., Velay, J.-L., & Besson, M. (2012). Twelve months of active musical training in 8- to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex, 24*(4), 956–967.

Coffey, E. B., Herholz, S. C., Chepesiuk, A. M., Baillet, S., & Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature Communications, 7*(11070), 2016. Doi:10.1038/ncomms11070

Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. *Clinical Neurophysiology, 112*(5), 758–767.

Engineer, N. D., Percaccio, C. R., Pandya, P. K., Moucha, R., et al. (2004). Environmental enrichment improves response strength, threshold, selectivity, and latency of auditory cortex neurons. *Journal of Neurophysiology, 92*(1), 73–82.

Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience, 6*(11), 1216–1223.

Gordon, K. A., Salloum, C., Toor, G. S., van Hoesel, R., & Papsin, B. C. (2012). Binaural interactions develop in the auditory brainstem of children who are deaf: Effects of place and level of bilateral electrical stimulation. *The Journal of Neuroscience, 32*(12), 4212–4223.

Gordon-Salant, S. (2014). Aging, hearing loss, and speech recognition: Stop shouting, I can't understand you. In A. N. Popper & R. R. Fay (Eds.), *Perspectives on auditory research* (pp. 211–228). New York: Springer Science + Business Media.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences, 15*(1), 3–10.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul H Brookes Publishing.

Hoff, E., Laursen, B., & Bridges, K. (2012). Measurement and model building in studying the influence of socioeconomic status on child development. In L. Mayes & M. Lewis (Eds.), *The Cambridge handbook of environment in human development* (pp. 590–606). Cambridge: Cambridge University Press.

Hornickel, J., & Kraus, N. (2013). Unstable representation of sound: A biological marker of dyslexia. *The Journal of Neuroscience, 33*(8), 3500–3504.

Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proceedings of the National Academy of Sciences of the U S A, 106*(31), 13022–13027.

Intartaglia, B., White-Schwoch, T., Meunier, C., Roman, S., et al. (2016). Native language shapes automatic neural processing of speech. *Neuropsychologia, 89*, 57–65.

Kilgard, M. P. (2012). Harnessing plasticity to understand learning and treat disease. *Trends in Neurosciences, 35*(12), 715–722.

Kilgard, M. P., & Merzenich, M. M. (1998). Cortical map reorganization enabled by nucleus basalis activity. *Science, 279*(5357), 1714–1718.

King, C., Warrier, C. M., Hayes, E., & Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neuroscience Letters, 319*(2), 111–115.

Kral, A., & Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends in Neurosciences, 35*(2), 111–122.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*(8), 599–605.

Kraus, N., & Nicol, T. (2005). Brainstem origins for cortical "what" and "where" pathways in the auditory system. *Trends in Neurosciences, 28*(4), 176–181.

Kraus, N., & Nicol, T. (2014). The cognitive auditory system. In R. R. Fay & A. N. Popper (Eds.), *Perspectives on auditory research* (pp. 299–319). Heidelberg: Springer.

Kraus, N., & White-Schwoch, T. (2015). Unraveling the biology of auditory learning: A cognitive-sensorimotor-reward framework. *Trends in Cognitive Sciences, 19*(11), 642–654.

Kraus, N., & White-Schwoch, T. (2016). Neurobiology of everyday communication: What have we learned from music? *The Neuroscientist.* Doi:10.1177/1073858416653593

Kraus, N., McGee, T., Carrell, T., King, C., Littman, T., & Nicol, T. (1994a). Discrimination of speech-like contrasts in the auditory thalamus and cortex. *The Journal of the Acoustical Society of America, 96*(5), 2758–2768.

Kraus, N., McGee, T., Littman, T., Nicol, T., & King, C. (1994b). Nonprimary auditory thalamic representation of acoustic change. *Journal of Neurophysiology, 72*(3), 1270–1277.

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., et al. (2014a). Auditory learning through active engagement with sound: Biological impact of community music lessons in at-risk children. *Frontiers in Neuroscience, 8*(351), 2016. Doi:10.3389/fnins.2014.00351

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., et al. (2014b). Music enrichment programs improve the neural encoding of speech in at-risk children. *The Journal of Neuroscience, 34* (36), 11913–11918.

Kraus, N., Thompson, E. C., Krizman, J., Cook, K., White-Schwoch, T., & LaBella, C. R. (2016). Auditory biological marker of concussion in children. *Nature, Scientific Reports, 6*(39009).

Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences of the U S A, 109*(20), 7877–7881.

Krizman, J., Skoe, E., & Kraus, N. (2016). Bilingual enhancements have no socioeconomic boundaries. *Developmental Science.* Doi:10.1111/desc.12347

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*(5044), 606–608.

Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *The Journal of Neuroscience, 29*(18), 5832–5840.

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin, 126*(2), 309–337.

Limb, C. J., & Braun, A. R. (2008). Neural substrates of spontaneous musical performance: An fMRI study of jazz improvisation. *PLoS ONE, 3*(2), e1679. Doi:10.1371/journal.pone.0001679

Loftus, W. C., Malmierca, M. S., Bishop, D. C., & Oliver, D. L. (2008). The cytoarchitecture of the inferior colliculus revisited: A common organization of the lateral cortex in rat and cat. *Neuroscience, 154*(1), 196–205.

Merzenich, M. M., Kaas, J., Wall, J., Nelson, R., et al. (1983). Topographic reorganization of somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation. *Neuroscience, 8*(1), 33–55.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the U S A, 104*(40), 15894–15898.

Musacchia, G., Strait, D., & Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and nonmusicians. *Hearing Research, 241*(1), 34–42.

Nieto-Diego, J., & Malmierca, M. S. (2016). Topographic distribution of stimulus-specific adaptation across auditory cortical fields in the anesthetized rat. *PLoS Biology, 14*(3), e1002397.

Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., et al. (2015). Family income, parental education and brain structure in children and adolescents. *Nature Neuroscience, 18*(5), 773–778.

Ozernov-Palchik, O., & Gaab, N. (2016). Tackling the "dyslexia paradox": Reading brain and behavior for early markers of developmental dyslexia. *Wiley Interdisciplinary Reviews: Cognitive Science, 7*(2), 156–176.

Pafundo, D. E., Nicholas, M. A., Zhang, R., & Kuhlman, S. J. (2016). Top-down-mediated facilitation in the visual cortex is gated by subcortical neuromodulation. *The Journal of Neuroscience, 36*(10), 2904–2914.

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009a). Musical experience limits the degradative effects of background noise on the neural processing of sound. *The Journal of Neuroscience, 29*(45), 14100–14107.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009b). Musician enhancement for speech-in-noise. *Ear and Hearing, 30*(6), 653–661.

Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012a). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging, 33*, 1483.

Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012b). Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Frontiers in Aging Neuroscience, 4*(30), 2016. Doi:10.3389/fnagi.2012.00030

Parbery-Clark, A., Strait, D. L., Hittner, E., & Kraus, N. (2013). Musical training enhances neural processing of binaural sounds. *The Journal of Neuroscience, 33*(42), 16741–16747.

Patel, A. D. (2008). *Music, language, and the brain*. New York, NY: Oxford University Press.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology, 2*(142), 2016. Doi:10.3389/fpsyg.2011.00142

Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *The Journal of Neuroscience, 26*(18), 4970–4982.

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. *The Journal of Neuroscience, 13*(1), 87–103.

Reed, A., Riley, J., Carraway, R., Carrasco, A., et al. (2011). Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron, 70*(1), 121–131.

Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment: Is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics, 31*(3), 509–527.

Russo, N., Skoe, E., Trommer, B., Nicol, T., et al. (2008). Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clinical Neurophysiology, 119*(8), 1720–1731.

Russo, N., Nicol, T., Trommer, B., Zecker, S., & Kraus, N. (2009). Brainstem transcription of speech is disrupted in children with autism spectrum disorders. *Developmental Science, 12*(4), 557–567.

Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G. (2013). Age-related cochlear synaptopathy: An early onset contributor to auditory functional decline. *The Journal of Neuroscience, 33*(34), 13686–13694.

Skoe, E., & Kraus, N. (2012). A little goes a long way: How the adult brain is shaped by musical training in childhood. *The Journal of Neuroscience, 32*(34), 11507–11510.

Skoe, E., Krizman, J., & Kraus, N. (2013). The impoverished brain: Disparities in maternal education affect the neural response to sound. *The Journal of Neuroscience, 33*(44), 17221–17231.

Slater, J., Tierney, A., & Kraus, N. (2013). At-risk elementary school children with one year of classroom music instruction are better at keeping a beat. *PLoS ONE, 8*(10), e77250.

Slater, J., Strait, D. L., Skoe, E., O'Connell, S., et al. (2014). Longitudinal effects of group music instruction on literacy skills in low-income children. *PLoS ONE, 9*(11), e113383.

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., et al. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research, 291*, 244–252.

Slee, S. J., & David, S. V. (2015). Rapid task-related plasticity of spectrotemporal receptive fields in the auditory midbrain. *The Journal of Neuroscience, 35*(38), 13090–13102.

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: Biological mechanisms. *Cerebral Cortex, 122*, 1890–1898.

Stevens, C., Lauinger, B., & Neville, H. (2009). Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: An event-related brain potential study. *Developmental Science, 12*(4), 634–646.

Strait, D. L., & Kraus, N. (2014). Biological impact of auditory expertise across the life span: Musicians as a model of auditory learning. *Hearing Research, 308*, 109–121.

Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency—effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience, 29*(3), 661–668.

Strait, D. L., Chan, K., Ashley, R., & Kraus, N. (2012a). Specialization among the specialized: Auditory brainstem function is tuned into timbre. *Cortex, 48*, 360–362.

Strait, D. L., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012b). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language, 123*, 191–201.

Strait, D. L., Parbery-Clark, A., O'Connell, S., & Kraus, N. (2013). Biological impact of preschool music classes on processing speech in noise. *Developmental Cognitive Neuroscience, 6*, 51–60.

Strait, D. L., O'Connell, S., Parbery-Clark, A., & Kraus, N. (2014). Musicians' enhanced neural differentiation of speech sounds arises early in life: Developmental evidence from ages three to thirty. *Cerebral Cortex, 24*(9), 2512–2521.

Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience, 5*(9), 721–728.

Tierney, A. T., & Kraus, N. (2014). Music training for the development of reading skills. In M. M. Merzenich, M. Nahum, & T. van Vleet (Eds.), *Applying brain plasticity to advance and recover human ability* (pp. 209–241). London: Elsevier.

Tierney, A., Krizman, J., Skoe, E., Johnston, K., & Kraus, N. (2013). High school music classes enhance the neural processing of speech. *Frontiers in Psychology, 4*(855), 2016. Doi:10.3389/fpsyg.2013.00855

Tierney, A., Krizman, J., & Kraus, N. (2015). Music training alters the course of adolescent auditory development. *Proceedings of the National Academy of Sciences of the U S A, 112*(32), 10062–10067.

van Rheede, J. J., Richards, B. A., & Akerman, C. J. (2015). Sensory-evoked spiking behavior emerges via an experience-dependent plasticity mechanism. *Neuron, 87*(5), 1050–1062.

Wang, B.-S., Sarnaik, R., & Cang, J. (2010). Critical period plasticity matches binocular orientation preference in the visual cortex. *Neuron, 65*(2), 246–256.

Warrier, C. M., Abrams, D. A., Nicol, T. G., & Kraus, N. (2011). Inferior colliculus contributions to phase encoding of stop consonants in an animal model. *Hearing Research, 282*(1–2), 108–118.

Webster, D. B. (1992). An overview of mammalian auditory pathways with an emphasis on humans. In D. B. Webster, A. N. Popper, & R. R. Fay (Eds.), *The mammalian auditory pathway: Neuroanatomy* (pp. 1–22). New York: Springer.

Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nature Reviews Neuroscience, 5*(4), 279–290.

Weiss, M. W., & Bidelman, G. M. (2015). Listening to the brainstem: Musicianship enhances intelligibility of subcortical representations for speech. *The Journal of Neuroscience, 35*(4), 1687–1691.

White-Schwoch, T., & Kraus, N. (2013). Physiologic discrimination of stop consonants relates to phonological skills in pre-readers: A biomarker for subsequent reading ability? *Frontiers in Human Neuroscience, 7*(899), 2016. Doi:10.3389/fnhum.2013.00899

White-Schwoch, T., Nicol, T., Warrier, C. M., Abrams, D. A., & Kraus, N. (in press). Individaul differences in human auditory processing: Insights from single-trial auditory midbrain activity in an animal model. *Cerebral Cortex*. Doi:10.1093/cercor/bhw29

White-Schwoch, T., Woodruff Carr, K., Anderson, S., Strait, D. L., & Kraus, N. (2013). Older adults benefit from music training early in life: Biological evidence for long-term training-driven plasticity. *The Journal of Neuroscience, 33*(45), 17667–17674.

White-Schwoch, T., Woodruff Carr, K., Thompson, E. C., Anderson, S., et al. (2015). Auditory processing in noise: A preschool biomarker for literacy. *PLoS Biology, 13*(7), e1002196.

Wible, B., Nicol, T., & Kraus, N. (2004). Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. *Biological Psychology, 67*(3), 299–317.

Wible, B., Nicol, T., & Kraus, N. (2005). Correlation between brainstem and cortical auditory processes in normal and language-impaired children. *Brain, 128*(2), 417–423.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*, 420–422.

Woodruff Carr, K., White-Schwoch, T., Tierney, A., Strait, D. L., & Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences of the U S A, 111*(40), 14559–14564.

Zendel, B. R., & Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychology and Aging, 27*(2), 410–417.

Zhu, X., Wang, F., Hu, H., Sun, X., et al. (2014). Environmental acoustic enrichment promotes recovery from developmentally degraded auditory cortical processing. *The Journal of Neuroscience, 34*(16), 5406–5415.