

# Chapter 4

## Short-Term Learning and Memory: Training and Perceptual Learning

Samuele Carcagno and Christopher J. Plack

**Abstract** The frequency-following response (FFR) is a sustained auditory-evoked potential that reflects the phase locking of neurons in the auditory brainstem to periodicities in the waveform of a sound. Studies have shown that short-term auditory training can improve the robustness and/or accuracy of this phase locking. FFR plasticity has been investigated using training tasks that are thought to involve some form of auditory temporal coding, including fundamental-frequency discrimination training, training to identify Mandarin lexical tones, and training to identify speech in noise. The results of these studies have shown that improvements in the trained task are often accompanied by FFR plasticity. This suggests that subcortical auditory processing is not hardwired but can be modified by training even in adulthood. The FFR has also been shown to change following auditory-cognitive training protocols in special populations of listeners who may have subcortical auditory processing deficits, such as children with language-based learning disabilities, elderly listeners, and listeners with sensorineural hearing loss. The results of these studies provide promising evidence that subcortical auditory plasticity could be harnessed to ameliorate auditory processing deficits. It has been hypothesized that this learning-induced subcortical plasticity may be guided by efferent cortical feedback; however, the mechanisms of FFR plasticity remain largely unclear.

**Keywords** Aging · Brainstem · Efferent system · Evoked potentials ·  $F_0$  discrimination · Learning disabilities · Lexical tones · Pitch · Plasticity · Speech-in-noise · Temporal coding

---

S. Carcagno (✉) · C.J. Plack  
Department of Psychology, Lancaster University, Lancaster LA1 4YF, UK  
e-mail: s.carcagno@lancaster.ac.uk

C.J. Plack  
e-mail: chris.plack@manchester.ac.uk

C.J. Plack  
Manchester Centre for Audiology and Deafness, University of Manchester,  
Manchester Academic Health Science Centre, Manchester M13 9PL, UK

## 4.1 Introduction

Sensory systems can change the way they process incoming external stimuli to extract from the environment information that is relevant for the organism. While some of these changes are fast and transient (e.g., adaptation of neural responses), perceptual learning reflects changes that generally occur over longer time scales and are long lasting (Goldstone et al. 2012). Perceptual learning is likely to play a crucial role in the acquisition of several skills, such as learning a foreign language (Lively et al. 1994), learning to play a musical instrument (Herholz and Zatorre 2012), or learning to detect tumors in X-rays (Sowden et al. 2000). Although the acquisition of some perceptual skills is thought to be limited to specific developmental time windows (Hensch 2004), mature perceptual systems retain a high degree of plasticity. This plasticity may be important not only to acquire new skills in adulthood but also to make effective use of novel or restored sensory stimulation following the introduction of sensory prostheses, such as cochlear implants (Fu and Galvin 2007; Moore and Shannon 2009), or to offset some of the deleterious effects of aging on perceptual systems (Anderson et al. 2013c; Bower et al. 2013).

The frequency-following response (FFR) has proved to be an invaluable tool to study perceptual learning and neural plasticity at the early stages of auditory processing. The initial interest in the FFR as a tool to study auditory plasticity at the level of the brainstem came from the finding that the FFRs of Mandarin speakers track the time-varying pitch of some Mandarin tones more accurately than the FFRs of English speakers (Krishnan et al. 2005; Krishnan and Gandour, Chap. 3). This finding could be interpreted as evidence of subcortical plasticity as a result of the extensive practice that Mandarin speakers have in the identification of Mandarin tones. However, the interpretation of such comparisons between experts and non-experts is equivocal because two nonrandomly selected groups of individuals may differ not only because of the amount of practice they have in a certain domain, but also for other reasons (Monaghan et al. 1998). For example, differences in FFR pitch-tracking ability could be potentially caused by genetic factors rather than by experience-related plasticity. This hypothesis is not a priori less plausible than the neuroplasticity hypothesis. Certain pitch processing abilities, such as the ability to detect a mistuned tone in a familiar melody, are highly heritable (Drayna et al. 2001). Also, the adoption of tone languages is associated with the frequency in the population of specific alleles of two genes related to brain growth, and this association is hard to explain by geographical or historical factors (Dediu and Ladd 2007). Genetic differences could explain both the behavioral and the neurophysiological advantage of Mandarin speakers in processing the dynamic pitch trajectory of Mandarin tones. Later studies of FFR differences between Mandarin and English speakers provided some evidence that these differences were due to experience, rather than to other factors, by showing that they were specific for tones with the same fundamental frequency ( $F_0$ ) contour as Mandarin tones (Xu et al. 2006). However, it is difficult to rule out genetic influences.

Musicians also have been found to have better FFR pitch tracking of tones than nonmusicians (Wong et al. 2007). This finding is consistent with the idea that extensive practice in discriminating and identifying tones leads to neuroplastic changes resulting in more accurate FFR pitch tracking. However, as for the comparison between Mandarin and English speakers, it is difficult to rule out the possibility that FFR differences between musicians and nonmusicians are due to factors other than neural plasticity.

Evidence for experience-dependent plasticity also has been provided by the results of short-term training studies in which FFR responses are recorded before and after a period of training in a tone discrimination or identification task. Unlike experiments comparing experts to nonexperts, short-term training studies can provide direct evidence of experience-related neuroplastic changes. These training experiments have additional advantages compared to experiments comparing experts to nonexperts, but they also have certain disadvantages. One advantage of training experiments is that they can determine whether or not FFR plasticity occurs in adulthood. Both Mandarin speakers and musicians usually start their experience in the discrimination and identification of Mandarin or musical tones during childhood. The superior FFR pitch tracking accuracy that these listeners show compared to English speakers without musical practice could be due to plastic changes that can happen only during a time-critical developmental window in childhood. Training experiments instead can test the hypothesis of adult neuroplasticity by directly training adults and testing for changes in the FFR. The question of whether the mature auditory system is malleable to experience-related changes is not only of theoretical interest. The potential for neuroplasticity in adulthood could pave the way for training interventions that are aimed at enhancing auditory skills in normal adults or remedy auditory deficits in special populations, such as elderly listeners with difficulties in hearing speech in noise or people with learning difficulties.

The disadvantages of training studies are that: (1) they are time consuming because usually participants have to attend multiple training sessions in addition to the sessions in which FFR responses are recorded; and (2) the amount of training that participants can get in a few sessions cannot match the often lifelong experience that Mandarin speakers and musicians have in processing tonal stimuli. As a consequence, the training effects elicited by short-term training studies may not be as large as the performance differences observed between expert and nonexperts.

The FFR is well suited for the study of auditory plasticity for several reasons. First of all, in the absence of auditory training, the FFR has good test-retest reliability (Song et al. 2011a, b). As a consequence, it is relatively easy to detect potential training-induced FFR changes. Another desirable feature of the FFR is that it can be recorded while participants are not paying attention to the stimuli (e.g., while watching a movie or even while sleeping). Under these conditions it is possible to exclude general changes in attention or higher-level cognitive processes as causes of FFR changes. Therefore, it is more likely that potential training-induced FFR changes reflect changes in sensory processing. Also, given the close correspondence between the FFR and certain stimulus features (Kraus et al. 2009), the interpretation of FFR changes

following training is generally easier than the interpretation of changes in the amplitude and/or latency of cortical event-related potential (ERP) components. In particular, given that the FFR reflects periodicities present in the stimulus, it is relatively straightforward to interpret FFR changes following training as enhancements in the fidelity with which these periodicities are represented in the brainstem. The FFR also offers practical advantages. FFR recordings with a good signal-to-noise ratio (SNR) can be obtained relatively quickly (e.g., for a single stimulus a good FFR response can be obtained with about 15 min or less of electroencephalogram recordings). The time needed to set up the electrodes with a single channel configuration (three electrodes: positive, negative, and ground), which has been commonly used in the literature, is also short. Recently, multichannel configurations have also been used for FFR recordings (Bharadwaj and Shinn-Cunningham 2014). While these multichannel setups require more time for electrode placement, with the use of multichannel denoising techniques they can achieve the same SNR as single channel configurations but with a reduced recording time. Therefore, these multichannel setups could be used to measure the time course of FFR changes within a recording session with a better time resolution than is possible with single channel setups.

The FFR also has certain limitations. Like all ERPs recorded at the scalp, it reflects the summed activity of many neurons and, therefore, can only give a relatively coarse picture of neural changes following training. On a positive note, however, the generators of the FFR have been localized to a relatively restricted region of the brainstem (Bidelman 2015). Like all ERPs, the FFR can be degraded by myogenic noise. Therefore, in studies of neuroplasticity care must be taken to avoid potential biases stemming from systematic differences in myogenic noise levels between sessions and/or between groups, depending on the study design. In particular, appropriate control groups should be used in between-subjects designs, or control stimuli in within-subjects designs.

Another limitation of the FFR is that due to frequency dependent delays of the traveling wave in the basilar membrane, the FFR to a low-frequency stimulus may reflect the activation of neurons with characteristic frequencies higher by an octave or more (Ananthanarayan and Durrant 1992; Dau 2003), while perceptual responses may result mainly from the activation of neurons with center frequencies closer to the stimulus frequency (Gockel et al. 2015). As a result, training-induced behavioral changes mediated by neurons with characteristic frequencies close to the stimulus frequency may not be reflected in the FFR. Finally, two practical limitations of the FFR in the study of neuroplasticity arise from the fact that the FFR can be readily recorded only in response to sounds of moderate or high intensity (sound levels greater than about 40–45 dB above perceptual threshold) and by tones with frequencies lower than about 2 kHz (Krishnan 2007). These limitations constrain the range of stimuli that can be used in studies of FFR plasticity.

The FFR is a sustained potential that reflects the phase locking of neural populations in the rostral brainstem to the envelope and/or fine structure of a sound waveform (Krishnan 2007). Therefore, the FFR reflects temporal information about a sound available at the level of the brainstem. Several auditory processes, such as sound localization (Grothe et al. 2010) and frequency coding (Moore 2008), are

thought to rely on precise temporal coding at the level of the brainstem. Thus, the FFR can be used as an index of temporal coding fidelity at the level of the brainstem to investigate whether and how temporal coding changes as a result of training-induced neural plasticity. According to temporal theories, the pitch of a sound is coded temporally in the peripheral auditory system (de Cheveigné 2005). Studies of FFR plasticity have commonly used pitch-discrimination or pitch-identification tasks for training. Accurate coding of temporal information is thought to play an important role also in more complex auditory skills, such as speech understanding in noise (Drennan et al. 2007; Moore 2008). Tasks involving speech understanding in noise have also been used in studies of FFR plasticity. Section 4.2 will summarize the main findings on the effects of short-term auditory training on behavioral and cortical electrophysiological responses, with a focus on pitch-discrimination training.

## 4.2 Effects of Short-Term Training on Behavioral and Cortical Responses

Many basic auditory skills, such as frequency and intensity discrimination (Wright and Zhang 2009) and complex auditory skills, such as speech understanding in noise (Engen 2012; Song et al. 2012), improve with practice. These improvements can result from changes at different processing levels. At the lowest level the response properties of early sensory neurons may change to sharpen the encoding of basic stimulus features (Gilbert et al. 2001). At a midlevel, the connections between sensory neurons and central decision areas may be re-weighted to enhance the internal representation of stimulus features that are crucial for the performance of the training task (Doshier and Lu 1999; Petrov et al. 2005). At the highest level, nonperceptual changes in decision processes, response biases, memory, and attentional resources can also influence the outcome of perceptual training (Amitay et al. 2014; Jones et al. 2015). Neurophysiological studies can help to pinpoint the source of training-induced changes in behavioral measures of performance by measuring neural responses at different processing levels before and after training. Caution needs to be exercised when drawing inferences from these studies. Changes of neural responses at higher processing stages may not reflect plasticity at those stages but rather reflect plasticity occurring downstream at lower processing stages. This argument also applies in reverse. The information flow does not simply proceed in a feedforward fashion from low-level sensory areas toward more central areas of the brain. Efferent feedback from central areas can modulate the responses of neurons in low-level sensory areas (Tsodyks and Gilbert 2004; Kraus and White-Schwoch 2015). Therefore, changes of neural responses in low-level sensory areas may not necessarily reflect plasticity at this level of processing (Watanabe and Sasaki 2015).

An important aspect of perceptual learning is its degree of specificity versus generalization. From a theoretical point of view, specificity of learning with respect to a given stimulus attribute or a given task can give important insights on the mechanisms underlying learning. If perceptual learning in a given task reflects a sharpening of tuning of early sensory neurons this learning should transfer to a different task in which performance is limited by the same noise in the sensory encoding phase. Task-specific learning, on the other hand, would indicate that learning affected later processing stages (Petrov et al. 2005). Specificity of learning with respect to a given characteristic of the trained stimulus has sometimes been used to infer the locus of training-induced plasticity. For example, in the visual system, specificity of learning to certain stimulus attributes, such as its retinal location or stimulus orientation, has been interpreted as a sign that learning is mediated by plasticity of primary sensory areas in which neural responses are best tuned to these basic perceptual attributes. However, this interpretation is equivocal (Mollon and Danilova 1996). Specificity of learning with respect to a low-level stimulus attribute such as retinal location, for example, may occur also as a result of a re-weighting of the connections between primary sensory areas and central decision areas (Doshier and Lu 1999; Petrov et al. 2005).

The degree of specificity of perceptual learning also has important implications from an applied point of view. Perceptual learning that is highly stimulus specific or task specific has little value in real-world scenarios in which organisms are confronted with stimuli that vary continuously along the trained perceptual dimension and appear in different contexts. For therapeutic or rehabilitative purposes, perceptual learning should generalize to a wide range of levels of the trained perceptual attribute and to different contexts. Some recent studies have started to investigate factors that promote the generalization of perceptual learning across different levels of the trained perceptual dimension and across presentation contexts. Deveau and Seitz (2014) have proposed that the use of a wide stimulus set, multisensory stimuli, reinforcement, and the engagement of attention are all factors that promote generalization of perceptual learning. The duration of training may also play an important role in the generalization of learning to untrained stimuli. Wright et al. (2010) have proposed that generalization of perceptual learning to untrained stimuli may lag behind stimulus-specific learning, and that increasing the duration of training may be an effective means to achieve greater generalization of learning to untrained stimuli.

One of the perceptual attributes that has been most widely studied in the context of auditory plasticity is pitch. Naturally occurring periodic sounds consist of several harmonically related frequencies, each an integer multiple of an  $F_0$ . The pitch of these sounds (i.e., the perceived tonal height) is commonly determined by their  $F_0$ , which usually corresponds to their envelope repetition rate. Pitch plays a crucial role in the perception of both speech and nonspeech sounds. In most languages the  $F_0$  conveys prosodic information as well as cues to the speaker's gender and emotional state (Rosen 1992; Skuk and Schweinberger 2014). In tone languages, the  $F_0$  is also used to convey lexical information; for example, in Mandarin Chinese two otherwise similar syllables take on a different meaning depending on the pitch

contour. In music the  $F_0$  plays a fundamental role in the composition of melodies and harmonies. The  $F_0$  also provides important cues for the perceptual grouping and segregation of auditory streams (Carlyon 2004). The ability to discriminate between sounds with a different  $F_0$  or a different  $F_0$  contour improves dramatically with practice. For this reason, as well as for the importance that pitch plays in hearing, tasks that rely on accurate pitch perception often have been used to study training-induced neural plasticity in the auditory system.

Several studies have found that training for multiple hours on an  $F_0$ -discrimination task leads to large improvements in performance on the task (see Wright and Zhang 2009 for a review). Micheyl et al. (2006) measured  $F_0$  discrimination thresholds in a group of musicians and a group of naive listeners. They found that musicians' thresholds were on average six times lower than the thresholds of naive listeners. However, the thresholds of nonmusicians decreased progressively as they trained on the task and became as low as those of musicians after 4–8 hours of training. These results indicate that short-term training effects on  $F_0$  discrimination are large and can match in magnitude the long-term effects of musical training after only a few hours of practice.

The behavioral improvements in pitch discrimination after training are often paralleled by changes of transient auditory-evoked potentials (AEPs) and auditory-evoked fields (AEFs), respectively derived from electroencephalographic and magnetoencephalographic recordings that reflect the activation of cortical areas in response to pitch-evoking sounds. The amplitude of the *N1m*, a sensory component of the AEFs peaking around 100 ms after the stimulus onset, has been found to increase after frequency-discrimination training (Menning et al. 2000), although the *N1*, the electric counterpart of the *N1m*, has not been found to change (Atienza et al. 2002; Bosnyak et al. 2004). A component of the AEPs whose amplitude has been found to increase after frequency-discrimination training is the *N1c*, a negative deflection occurring around 150 ms after stimulus onset at temporal electrode sites (Bosnyak et al. 2004). Several studies have also found increases in the amplitude of the *P2* component of the AEPs, a positive deflection peaking around 180 ms after stimulus onset (Bosnyak et al. 2004; Carcagno and Plack 2011a). The functional significance of these *P2* changes has been a matter of debate because increases in the amplitude of the *P2* have also been found following passive exposure to sounds rather than active learning (Sheehan et al. 2005). In a study by Carcagno and Plack (2011a), however, these *P2* amplitude changes extended also to untrained stimuli, suggesting that stimulus exposure per se could not explain increases of the *P2*. Recently, it has been suggested that *P2* changes are associated with some aspect of the training procedure itself rather than with perceptual learning (Tremblay et al. 2014). The mismatch negativity (MMN), a difference wave derived by subtracting the AEPs to a rarely presented “deviant” tone from the AEPs to a frequently presented “standard” tone, has also been found to increase in amplitude after training on pitch-related tasks (Atienza et al. 2002; Tong et al. 2009). Carcagno and Plack (2011a), however, did not find MMN changes after training participants on an  $F_0$ -discrimination task despite the fact that the authors demonstrated that the measured MMNs had an adequate sensitivity to detect equivalent changes in

performance produced by changes to the stimuli rather than by training. The reasons for this discrepancy remain unclear but may be related to differences in the task (deviance detection versus  $F_0$  discrimination) used for the training, suggesting that MMN-amplitude increases may reflect changes of higher-level melodic processing functions rather than low-level sensory discrimination processes.

Overall, AEP and AEF studies indicate that short-term training affects the cortical processing of pitch-evoking sounds at several levels of processing. The functional significance of these changes, however, is not yet well understood. Given that AEPs and AEFs in these studies were recorded while participants were passively listening to the stimuli, rather than being actively engaged in a task, it is likely that these AEP/AEF changes reflect changes of low-level sensory processes. One possibility is that they reflect reorganization of cortical tonotopic maps. Studies of tonotopic map plasticity following perceptual discrimination training in non-human animals are few, and the evidence for such tonotopic reorganization is mixed (Recanzone et al. 1993; Brown et al. 2004). Recent results suggest that an expanded cortical representation of the trained stimulus accelerates learning in the initial stages of training. However, this map expansion is subsequently renormalized and is not necessary for maintaining good performance in a learned discrimination task (Reed et al. 2011). It is also possible that training-induced changes in cortical AEP/AEF responses reflect, at least in part, training-induced changes at more peripheral levels of processing.

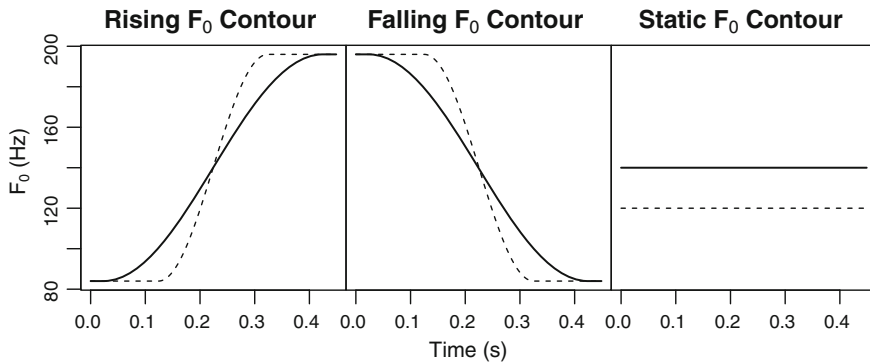
### 4.3 Effects of Short-Term Training on FFR Responses

Until about a decade ago, experience-related auditory plasticity was studied mainly using cortical responses. Following the findings that Mandarin speakers (Krishnan et al. 2005) as well as musicians (Wong et al. 2007) show enhanced subcortical encoding of Mandarin tones, suggesting that experience-related plasticity may also occur in brainstem structures, several studies have investigated the effects of training on the subcortical encoding of sounds using the FFR. Several training tasks have been used to investigate FFR plasticity, including pitch and pitch-contour discrimination, identification of lexical tones, and speech-in-noise comprehension. This section will summarize the FFR changes that have been found after training on these tasks in normal-hearing listeners as well as in special populations of listeners with hearing difficulties.

#### 4.3.1 *Effects of Pitch-Discrimination Training*

As mentioned in Sect. 4.2, short-term training on a pitch discrimination task can lead to large improvements in pitch discrimination after only a few hours of training. Carcagno and Plack (2011b) trained three groups of listeners on an  $F_0$



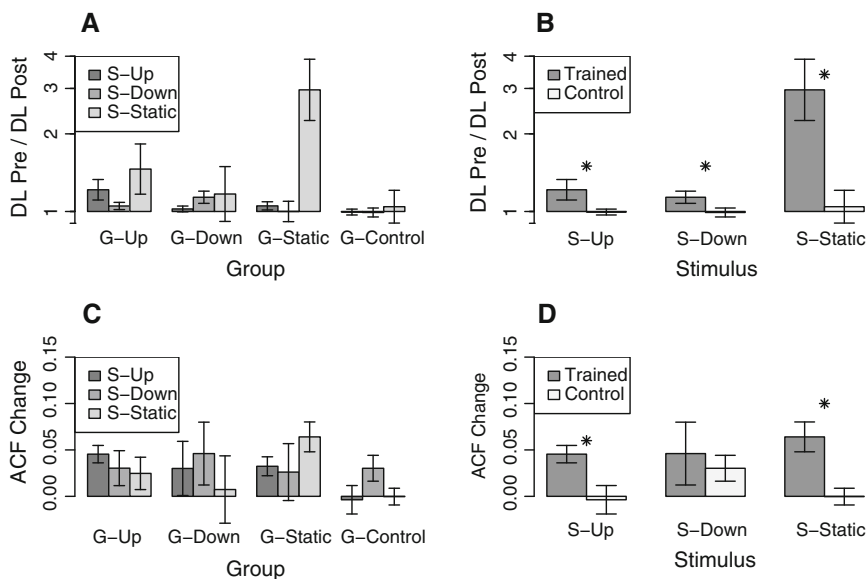


**Fig. 4.1**  $F_0$  contours of the stimuli used in the study by Carcagno and Plack (2011b). The *solid lines* represent the  $F_0$  contours of the standard stimuli. The *dashed lines* show examples of the comparison stimuli that had to be discriminated from the standard stimuli. For the stimuli with a dynamic  $F_0$  contour (first two panels) the comparison stimuli had a higher FM rate. For the stimulus with a static  $F_0$  (right panel) the comparison stimuli had a lower  $F_0$

discrimination task for 10 hours and compared the difference between their pre-training and post-training FFR responses to those of an untrained control group. Each group was trained with one of three different stimuli (depicted in Fig. 4.1).

All three stimuli consisted of harmonic complex tones band-pass filtered between 2 and 3 kHz, so as to include only unresolved harmonics (see Plack and Oxenham 2005 for a definition of resolvability). The FFR reflects temporal information, thus, improvements on a temporal pitch-encoding mechanism are likely to be reflected in the FFR (improvements on a place pitch-encoding mechanism, based on cochlear frequency selectivity, would not necessarily be reflected in the FFR). The use of complex tones consisting exclusively of unresolved harmonics forced listeners to use a temporal pitch-encoding mechanism to perform the task because these tones do not contain useful spectral information to extract their pitch. Two of the stimuli (S-Up and S-Down) had a dynamic  $F_0$  contour, while the third stimulus (S-Static) had a static  $F_0$  contour. For the dynamic stimuli, the task was to detect a difference in the rate of change of  $F_0$ . For the static stimulus, the task was to detect an  $F_0$  difference.

Before and after training the behavioral discrimination thresholds as well as the FFR responses to each stimulus were measured for all listeners (Figs. 4.2A, B). For each trained group, the  $F_0$ -discrimination performance for the trained stimulus improved more than for the control group, indicating that training was effective in eliciting  $F_0$  discrimination learning. For listeners trained with a dynamic  $F_0$  contour, these performance improvements were partly specific to the trajectory (rising versus falling) of the  $F_0$  contour with which they were trained. Collapsing the data of the stimuli with a dynamic  $F_0$  contour (S-Up and S-Down) together and comparing them to the data from the stimulus with an unmodulated  $F_0$  contour (S-Static), it was also apparent that threshold improvements were partly specific to the modulation (dynamic versus static) of the  $F_0$  contour used during training.



**Fig. 4.2** Changes in the behavioral and FFR measures in the study of Carcagno and Plack (2011b). (A) Ratio of the behavioral discrimination thresholds before and after training. *G-Up*, *G-Down*, *G-Static*, and *G-Control* labels denote respective training groups (e.g., *G-Up* trained on the *S-Up* stimulus). Values >1 indicate improvements. (B) Plot as in A, but for a set of planned contrasts comparing changes for each trained group on the trained stimulus to changes in the control group for the same stimulus. (C) Difference of the autocorrelation function (ACF) value at the time lags corresponding to the inverse of the  $F_0$  of the stimuli between post-training and pre-training FFR assessment sessions. (D) Plot as in C, but for a set of planned contrasts comparing changes for each trained group on the trained stimulus to changes in the control group for the same stimulus. The asterisks in B and D denote statistically significant differences ( $p < 0.05$ ). (Reproduced with permission from Carcagno and Plack 2011b)

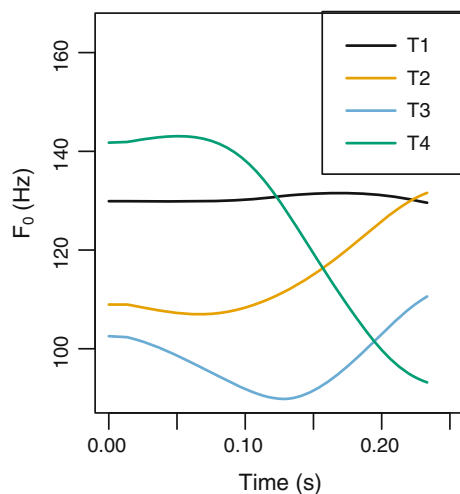
The robustness of FFR phase locking to the sound envelope of each stimulus before and after training was assessed by averaging the value of the autocorrelation function (ACF) at the time lag corresponding to the inverse of the stimulus  $F_0$  in seven time windows of 64-ms length. Figure 4.2 (C, D) shows the difference between the post-training and pre-training average value of the ACF for each stimulus and group. The ACF value for the trained stimulus increased significantly more for participants trained with the *S-Up* and *S-Static* stimuli than for participants in the control group. For participants trained with the *S-Down* stimulus, the ACF change for the trained stimulus was not significantly larger than for control participants. As for threshold changes, ACF changes were partly specific to the modulation (dynamic versus static) of the  $F_0$  contour of the trained stimulus, while the specificity of ACF changes to the  $F_0$  trajectory (rising versus falling) was not statistically significant. The behavioral threshold improvements were found to correlate with the ACF changes for participants trained with the *S-Up* and *S-Static* stimuli, while the correlation for participants trained with the *S-Down* stimulus was not significant.

Overall, the results of this study indicate that the robustness of subcortical temporal encoding can improve as a result of short-term  $F_0$  discrimination training. The correlation between FFR and threshold changes indicates that perceptual improvements are associated with FFR changes. However, it should be kept in mind that these correlations were relatively small, accounting for about 14–20% of the variance in threshold improvements for participants trained with the S-Up and S-Static stimuli. These data suggest that although the subcortical changes in temporal encoding measured by the FFR make a significant contribution to improvements in pitch discrimination after training, these behavioral improvements may depend in large part either on subcortical changes that are not captured by the FFR or by changes occurring at higher levels of processing.

### 4.3.2 Effects of Training on the Identification of Lexical Tones

Several studies have used tasks in which participants are presented with pseudo-words consisting of syllables onto which the  $F_0$  contour of a Mandarin lexical tone is superimposed, and they are trained to associate these pseudo-words with drawings of English nouns. In order to successfully perform this task, participants must be able to follow the dynamic  $F_0$  contour of each Mandarin tone used and discriminate it from the others. Song et al. (2008) trained 23 English speakers in a lexical tone identification task for eight sessions each lasting for about 30 min. The FFR responses to lexical tones were recorded before and after training. Mandarin uses four lexical tones: *T1* (high-level), *T2* (rising), *T3* (falling-rising), and *T4* (falling), with different  $F_0$  contours (depicted in Fig. 4.3).

**Fig. 4.3** Prototypical  $F_0$  contours of the four Mandarin lexical tones. The  $F_0$  contours were derived from the synthesized speech monosyllable /yi/ following the procedures used by Krishnan et al. (2005). (Data courtesy of Ananthanarayan Krishnan)



Three of these tones (T1, T2, and T3) were used during training. The percentage of correct identifications across all three tones improved from 22% after the first training session to 89% after the last training session. FFR pitch-tracking accuracy, assessed by measuring the distance between the  $F_0$  trajectories extracted from spectrograms of the FFR and stimuli waveforms, improved after training for the T3 tone. This improvement was paralleled by an increase in SNR measured by comparing the power at the  $F_0$  during the presentation of the stimulus to the power at the same frequency in the pre-stimulus baseline window. No significant post-training changes were found for the T1 and T2 tones. Although English is not a tonal language,  $F_0$  modulations similar to T1 and T2 occur frequently in English at the syllable level as intonational markers, while  $F_0$  modulations following the T3 pattern do not occur in English at the syllable level. Song et al. (2008) argued that the familiarity of English speakers with the T1 and T2 pitch contours may explain why FFR pitch tracking did not appear to improve for these stimuli and improved only for the T3 tone, which was the least familiar to the participants. Other studies, however, have subsequently found training effects also for these tones.

Chandrasekaran et al. (2012) investigated the impact of individual differences in inferior colliculus (IC) activity on lexical tone learning assessed behaviorally and electrophysiologically via the FFR. The authors trained 18 English speakers on a lexical tone identification task for nine sessions. Before training fMRI scans of the IC were collected in response to repeated lexical tones (e.g., T1-T1-T1-T1, tone repeat condition) or alternating lexical tones (e.g., T1-T3-T4-T2, tone change condition). These scans were used to divide the participants into a repetition suppression (RS) group, showing less activation in the tone repeat condition than in the tone change condition, and a repetition enhancement (RE) group, showing the reverse pattern. RS is generally associated with improved neural encoding of the repeated signal and, according to neural sharpening models (Grill-Spector et al. 2006), reflects a sparser and more efficient representation of the signal (fewer neural units are needed to encode it). Before and after training, FFR amplitude and FFR pitch tracking accuracy (based on stimulus-to-response correlations of  $F_0$  trajectories derived by autocorrelograms) were measured for tone T2. FFR pitch tracking was measured separately for the initial portion of the stimulus (in which the  $F_0$  was nearly static) and for the second part of the stimulus (in which the  $F_0$  was rising).

Activation of the IC (RS versus RE) was predictive of pre-training FFR amplitude, with participants in the RS group showing smaller FFR amplitudes (possibly reflecting greater neural efficiency) than participants in the RE group. The RS participants also had better FFR pitch-tracking accuracy than the RE participants for the dynamic portion of the stimulus. This is consistent with the idea that an efficient representation of the stimulus is associated with better encoding of the stimulus.

Performance in the lexical tone-identification task was initially similar for the two groups but, from the third session onward, participants of the RS group showed greater improvements in the task than participants of the RE group. After training, FFR pitch-tracking accuracy improved significantly for both groups on the dynamic-pitch portion of the stimulus, while no training-related changes were seen

for the static- $F_0$  portion of the stimulus. The FFR amplitude, on the other hand, decreased significantly for the RS group, but not for the RE group. In addition to providing further evidence of greater post-training accuracy of FFR pitch tracking, these results indicate that pre-training FFR amplitude may be an index of stimulus encoding efficiency at the level of the brainstem and may be predictive of learning success in a lexical tone-identification task.

A number of studies have shown that the FFR is sensitive to the probability of presentation of a stimulus (see Escera, Chap. 5). Skoe et al. (2014) investigated how these probability effects interact with the effects of short-term training on the accuracy of  $F_0$  encoding. The authors trained 12 participants for nine sessions using a lexical tone identification task. FFR pitch-tracking accuracy was assessed by taking the distance between the  $F_0$  contours of the FFR and stimulus waveforms, which were derived from autocorrelograms. FFR pitch-tracking accuracy was measured before and after training and compared to that of a control group consisting of 13 participants who did not receive training. Stimuli (tones T1 and T4) were presented in an oddball paradigm (80% standards; 20% deviants) to investigate the effects of stimulus probability. Before training, FFR pitch tracking was more accurate for frequently presented standard sounds than for deviants. After training, collapsing responses across standard and deviant tones, FFR pitch-tracking accuracy improved for the trained group but not for the control group. Additionally, the effect of stimulus probability for tone T1 reversed after training: post-training FFR pitch-tracking accuracy was greater when T1 tones were presented as deviants compared to when they were presented as standard tones. In addition to providing further evidence that FFR pitch-tracking accuracy can be improved by training, these results suggest that the effects of training interact in a complex way with contextual effects of stimulus probability.

### ***4.3.3 Effects of Training to Understand Speech in Noise***

FFR plasticity has also been investigated using general speech-in-noise training protocols. Song et al. (2012) tested 28 participants that trained at home for 20 sessions (approximately 30 min per session) using the Listening and Communication Enhancement (LACE<sup>®</sup>; Neurotone, Inc., Redwood City, CA) program and 32 participants who did not receive training. The LACE<sup>®</sup> program included tasks related to comprehension of degraded speech, cognitive skills, and communication strategies. Before and after training, FFR responses to the syllable /da/ were recorded in quiet and in a six-talker babble noise. Participants were also tested for their ability to hear speech in noise using the Quick Speech-in-Noise Test (QuickSIN<sup>™</sup>; Etymotic Research, Elk Grove Village, IL) and the Hearing-in-Noise Test (HINT; Biologic Systems Corp., Mundelein, IL). A subset of the participants was re-tested approximately six months after the end of the program to check for retention of training effects. The FFR was analyzed separately for the formant transition region (20–60 ms) and the steady-state vowel

region (60–180 ms). The average of the FFR response measured from FFT spectra at frequencies corresponding to the first two harmonics ( $F_0$  and H2) of the stimulus was also analyzed separately from the average of the FFR response at frequencies corresponding to harmonics 3 to 10 of the stimulus.

After training, trained participants showed larger improvements than control participants on the QuickSIN<sup>TM</sup> and HINT tests. These behavioral improvements were paralleled by a larger increase of the FFR responses in noise at frequencies corresponding to the first two harmonics of the stimulus in the trained group compared to the control group. This effect was present for both the steady state and the formant transition region, although it was significantly greater for the latter. No significant differences between the trained and control group were found for the FFR responses to the syllable in quiet. Both the behavioral and the neurophysiological improvements were in large part retained six months after the training. Another interesting finding was that the strength of  $F_0$  and H2 encoding measured by the FFR at pretest predicted the improvement in some of the speech-in-noise measures; participants who had better FFR encoding of  $F_0$  and H2 at pretest showed the largest amount of learning. The results of this study are important because, in addition to showing that a general speech-in-noise training protocol can improve subcortical encoding of temporal information, they also provide evidence of long-term retention of both behavioral and FFR training effects.

#### ***4.3.4 Studies on Special Populations***

The finding that subcortical auditory processing is not static but can be manipulated by training suggests that sensory deficits caused by degraded subcortical temporal processing may be partly remedied by training. A number of studies have tested this hypothesis by measuring subcortical responses before and after a period of auditory training in special populations of listeners with hearing difficulties.

Children with language-based learning problems (e.g., dyslexia) often show concomitant auditory perception problems. Russo et al. (2005) trained nine children with learning disabilities for 35 to 40 one-hour sessions with *Earobics*<sup>®</sup> (Cognitive Concepts, Inc., Evanston, IL), a commercial auditory training program. The training consisted of interactive computer games that tapped into phonological awareness, auditory skills, and language processing skills. Auditory stimuli were presented both in quiet and with background noise. Brainstem and cortical responses to the syllable /da/, as well as measures of auditory and cognitive processing, were evaluated before and after training. A control group, consisting of five children with learning disabilities and five normal children, was also evaluated on these measures across a time period equivalent to the training period but did not undergo training. After training, the correlation between the FFR component of the response recorded in quiet and the FFR component of the response recorded in noise increased significantly more for the experimental group compared to the control group. This

increased inter-response correlation reflected mainly an improved phase locking of the response in noise. Children in the training group also showed improvements in several behavioral tests, including the Incomplete Words, Auditory Processing, and Sentences-in-Noise tests. It is unclear how many of the children tested by Russo et al. (2005) specifically had a language-based learning disability. However, language-based learning disabilities constitute a sizable portion of learning disabilities, and given the extensive literature linking these disabilities to subcortical timing deficits (for reviews see Chandrasekaran and Kraus 2012; Reetzke, Xie, and Chandrasekaran, Chap. 10), the results of Russo et al. (2005) are a promising line of work for the development of training interventions for children with these disabilities.

Additional evidence for the plasticity of subcortical auditory processing in children with language-based learning disabilities comes from a study of Hornickel et al. (2012). In this study, the FFR to the syllables /ba/, /da/, and /ga/ was recorded at the beginning and at the end of one academic year in 38 children (8–14 years old) diagnosed with dyslexia. Reading and phonological awareness tests were also administered at each time point. During the academic year, 19 children assigned to the treatment group wore *FM systems* listening devices in the classroom, while the remaining 19 children in the control group did not. FM systems are radio systems that broadcast the voice of a speaker wearing a microphone directly to the ears of a listener wearing a receiver, thus bypassing background noise and enhancing the acoustic clarity of the speaker's voice. At the end of the academic year, children in the treatment group had improved scores on the reading and phonological awareness tests, while children in the control group did not show significant improvements in these tests. At the end of the year, the consistency of the FFR, as measured by the correlation between the average response obtained during the first and the second half of the session, improved significantly in the formant transition region of the syllables (7–60 ms) for the children in the treatment group. No significant changes in FFR response consistency were found for children in the control group. Interestingly, the greatest improvements in FFR consistency were observed in a subset of children who showed the greatest improvements in the phonological awareness tests. Overall, these results suggest that the use of FM systems listening devices may improve both reading abilities and subcortical responses to sounds in children with dyslexia. The authors of the study hypothesize that these changes are mediated by the improved clarity of the acoustic signal provided by the FM system. The increased acoustic clarity of the acoustic signal may strengthen sound-to-meaning relationships and, through the action of the efferent system, may lead to a fine-tuning of the temporal encoding of sounds at a subcortical level.

Elderly listeners represent another population of listeners that might benefit from auditory perceptual training. Recent studies indicate that age-related visual processing deficits can be partially offset by perceptual training (Andersen et al. 2010; Polat et al. 2012). Similarly, reversals of age-related deficits have been observed in the cortical responses of rats after auditory perceptual training (de Villers-Sidani et al. 2010). Subcortical auditory temporal processing degrades as a result of aging. FFR phase locking to pure tones (Clinard et al. 2010; Marmel et al. 2013) and

musical dyads (Bones and Plack 2015) has been found to decline as a function of age. Also, the timing of peaks in the subcortical response to syllables (Anderson et al. 2012; Clinard and Tremblay 2013) has been found to be delayed as a result of aging (see Anderson, Chap. 11).

Anderson et al. (2013c) investigated whether a home-based cognitive training program could offset some of these age-related deficits. The authors trained 35 elderly participants (ages 55–70) for 8 weeks (40 hours of training at home) with *Brain Fitness*<sup>TM</sup> (Posit Science Corporation, San Francisco, CA), an auditory-based cognitive training program, while 32 control participants in the same age range watched a series of educational DVDs (about art, science, history and other topics) and completed multiple-choice tests on the topics of the DVDs. The *Brain Fitness*<sup>TM</sup> program contains exercises in which the formant transition region of speech syllables is adaptively stretched or compressed. The exercises are presented at different levels (isolated syllables, syllables within words, sentences, and stories). Before and after training, brainstem responses to the syllable /da/ were recorded, as well as measures of speech perception in noise (QuickSIN<sup>TM</sup> test), auditory short-term memory, and processing speed. After an initial stop burst of 5 ms, the syllable had a constant  $F_0$  of 100 Hz with peaks occurring every 10 ms. After training, the timing of the FFR peaks in the formant transition region of the syllable in noise was significantly earlier for the trained group than for the control group. Across the overall length of the stimulus presented in noise, the FFR inter-peak variability, measured as the standard deviation of latency differences between adjacent peaks, was also significantly reduced after training in the experimental group compared to the control group. These neurophysiological changes were paralleled by significant improvements in all behavioral measures (QuickSIN<sup>TM</sup> test, auditory short-term memory, and speed processing) in the trained group compared to the control group.

Anderson et al. (2014) re-tested 30 of the 35 participants in the auditory training group and all 32 participants of the active control group in a follow-up study six months after the training to determine whether or not behavioral and neurophysiological changes would be retained in the long term. The results of this follow-up study revealed that subcortical improvements in response peak timing were largely maintained after six months. Behaviorally, improvements in speed processing were also maintained, but gains in auditory short-term memory and speech reception thresholds were not. The dissociation between retention of subcortical peak timing changes and changes in speech reception thresholds suggests that the two were not directly related in this study. Improvements in speech reception thresholds immediately after the training may have instead resulted from cognitive components of the training. In sum, the results of the studies of Anderson et al. (2013c, 2014) indicate that perceptual learning can lead to subcortical plasticity also in elderly listeners and can ameliorate some age-related perceptual auditory deficits. However, more work needs to be done to understand the relationship between neurophysiological and behavioral changes following training and how the benefits of training can be retained over time.



Additional evidence of subcortical plasticity following training in elderly listeners comes from a study by Anderson et al. (2013b) in which two groups of elderly participants, one with normal hearing and one with mild hearing loss, were trained for eight weeks with *Brain Fitness<sup>TM</sup>*, and their performance changes in a series of behavioral and electrophysiological tests were compared to those of age and audiometrically matched active control groups as in the study by Anderson et al. (2013c). Because listeners with hearing loss have abnormally enlarged responses to sound envelopes (Anderson et al. 2013a), this study focused on the relative representation of envelope and fine structure in the FFR. The pre-training results confirmed the presence of enlarged FFR envelope responses in the group of listeners with hearing loss. These enlarged FFR envelope responses were reduced in the group of hearing-impaired participants after training compared to the age and audiometrically matched control group. No significant changes occurred for the normal-hearing participants. This reduction of the FFR envelope response was accompanied by small improvements in speech-reception thresholds. The effect of training on the FFR in this study, a reduction of the FFR envelope response, is opposite to the enhancement of the FFR envelope response found in two previous studies of short-term FFR plasticity (Carcagno and Plack 2011b; Song et al. 2012). The cause of this difference is unclear. It may be related to the different populations of listeners employed in these studies (young normal-hearing versus elderly hearing-impaired listeners) or to differences in the training tasks and stimuli employed.

### 4.3.5 Summary of Empirical Findings on FFR Plasticity

Table 4.1 presents an overview of the short-term training studies on FFR plasticity that have been conducted to date. For each study the table lists the stimuli and training task used and specifies whether for each stimulus a significant training effect on the FFR was found on at least one of the measures used to quantify the FFR changes. It should be kept in mind that this is only a qualitative and crude way to summarize FFR training effects, but a full-blown meta-analytic treatment is beyond the scope of this chapter. The table also includes the population of listeners trained in each study and whether their results were compared to those of a control group or not.

Overall, the studies reviewed above indicate that short-term training on pitch discrimination or pitch identification leads to increased FFR pitch-tracking accuracy. The term “FFR pitch tracking” is used in this chapter to indicate the fact that FFR periodicities follow periodicities present in the stimulus. However, it should be noted that the FFR is not a direct measure of the subcortical representation of pitch; rather, it reflects subcortical timing information that may be used by the auditory system to encode pitch (Gockel et al. 2011). It is currently unclear how specific

**Table 4.1** Summary of short-term training FFR studies

Study	Stimulus	Task	Treatment group	Control group	Follow-up
Carcagno and Plack (2011b)	Rising F <sub>0</sub> -Sig. Falling F <sub>0</sub> -NS Static F <sub>0</sub> -Sig.	F <sub>0</sub> discrimination	Young adults (N = 27; 9 per training group)	Young adults (N = 12) Passive	No
Song et al. (2008)	T1-NS T2-NS T3-Sig.	Lexical tone identification	Young adults (N = 23)	No	No
Chandrasekaran et al. (2012)	T2 (rising segment)-Sig. T2 (static segment)-NS	Lexical tone identification	Young adults (N = 18)	No	No
Hornickel et al. (2012)	/ba/-/da/-/ga/- Sig.	Wearing FM assisting listening devices	Children with dyslexia (N = 19)	Children with dyslexia (N = 19)	No
Skoe et al. (2014)	T1-Sig. T4-Sig.	Lexical tone identification	Young adults (N = 12)	Young adults (N = 13) Passive	No
Song et al. (2012)	/da/-Sig.	Speech-in-noise comprehension	Young adults (N = 28)	Young adults (N = 32) Passive	6 months Maintained
Russo et al. (2005)	/da/-Sig.	Phonological, auditory, and language processing: <i>Earobics</i> <sup>®</sup>	Children with learning disabilities (N = 9)	Children with (N = 5) and without (N = 5) learning disabilities	No
Anderson et al. (2013c)	/da/-Sig.	Auditory cognitive training: <i>Brain Fitness</i> <sup>TM</sup>	Elderly adults (N = 35)	Elderly adults (N = 32) Active	6 months Largely maintained
Anderson et al. (2013b)	/da/-Sig.	Auditory cognitive training: <i>Brain Fitness</i> <sup>TM</sup>	Elderly adults with (N = 15) or without (N = 14) hearing loss	Elderly adults with (N = 15) or without (N = 14) hearing loss Active	No

*Stimulus* column: stimuli used and whether training had a significant (Sig.) or not significant (NS) effect for at least one of the measures used to quantify FFR changes. *Treatment group*: specifies the population tested and the number of participants. *Control group*: specifies whether a control group was used or not and, if applicable, the number of participants included in the control group and whether the participants in the control group were engaged in an alternative training task (*Active*) or not (*Passive*). *Follow-up*: indicates the time at which any follow-up test was performed and whether FFR changes measured immediately after training were maintained at the follow-up test

these improvements in FFR pitch-tracking accuracy are with respect to the trained  $F_0$  contour. The only study that investigated this issue (Carcagno and Plack 2011b) found that improvements were specific to the presence/absence of a modulation in the  $F_0$  contour, but they did not find evidence of specificity with respect to the trajectory (rising or falling) of a modulated  $F_0$  contour. It also remains unclear whether certain  $F_0$  contours are more likely to elicit FFR training effects compared to others. Stimuli with dynamically modulated  $F_0$  contours seem more likely to elicit FFR training effects compared to stimuli with a static  $F_0$  contour, possibly as a result of the fact that tracking a dynamically changing  $F_0$  is more challenging and susceptible to tracking errors than encoding a static  $F_0$ . There is no clear evidence of larger FFR training effects for a particular  $F_0$  contour among stimuli with dynamically modulated  $F_0$  contours. For some Mandarin tones training effects were found in some studies (T2: Chandrasekaran et al. 2012; T1: Skoe et al. 2014) but not all (Song et al. 2008). This apparent discrepancy may reflect methodological differences (e.g., differences in the measures used to summarize FFR training effects), but it should also be kept in mind that in single studies with relatively modest sample sizes effect sizes may be either overestimated or underestimated. Only a meta analysis combining the results of several studies could clarify whether FFR training effects for these tones are smaller than for other Mandarin tones.

Speech perception training tasks seem to have positive effects on subcortical temporal encoding, resulting in more robust encoding of stimulus-related periodicities. These training benefits appear to be greater for stimuli presented in noise. Again this may reflect the fact that the encoding of stimulus features in noise is a more challenging task for the auditory system and more susceptible to encoding errors that may be remedied by training.

With the exception of the studies of Anderson et al. (2013b, c), none of the studies reviewed above used an active control group. This makes it difficult to determine whether some of the measured changes reflect practice in the specific training task or general effects of auditory training. Some of the studies did not have a control group at all (Song et al. 2008; Chandrasekaran et al. 2012), leaving open the possibility that the measured changes were not a direct consequence of training. Furthermore, in the study of Hornickel et al. (2012), even though a control group was included, FFR changes in the treatment group were not directly compared to FFR changes in the control group. Instead, a statistically significant FFR change in the treatment group, and a lack of statistically significant FFR change in the control group were taken as evidence of FFR plasticity. Unfortunately, this kind of comparison, although suggestive of such an effect, does not warrant such a conclusion (Nieuwenhuis et al. 2011). Given that the measured changes (e.g., accuracy of FFR pitch tracking) are closely linked to the training task (e.g.,  $F_0$  discrimination) it may seem unlikely that they would be caused by factors other than the training itself. However, FFR pitch tracking may improve not only as a result of changes in specific temporal encoding mechanisms but also as a result of a generic reduction in physiological noise, for example, due to the participant being more relaxed in later test sessions. Therefore, it is important that future studies of FFR training-induced changes use appropriate controls to identify the specific causes of FFR changes.

#### 4.4 Mechanisms of Auditory Brainstem Plasticity in Short-Term Learning

There is currently no comprehensive mechanistic model of auditory brainstem plasticity induced by perceptual learning. This is perhaps not surprising because, besides the fact that interest in subcortical plasticity is relatively new, such models are likely to involve complex interactions between brainstem nuclei and cortical structures via the efferent system.

In vitro studies in brain slices indicate that cells from auditory brainstem nuclei display synaptic plasticity in the form of long-term potentiation and long-term depression (reviewed by Tzounopoulos and Kraus 2009). While to the best knowledge of the authors there have been no neurophysiological studies of subcortical plasticity induced by perceptual-discrimination training in nonhuman animals, several studies have investigated subcortical auditory plasticity induced by associative learning (for a discussion of differences between perceptual and associative learning see Weinberger 2008). These studies have shown that after auditory fear conditioning with a tone, the best frequency of IC neurons in the big brown bat (*Eptesicus fuscus*) shifted toward the frequency of the conditioned tone. These best-frequency shifts are short term, lasting only about 3–3.5 hours after the association phase (Suga and Ma 2003). Inactivation of the auditory cortex during the association phase blocks these best-frequency shifts, indicating that plasticity is mediated by corticofugal modulation via the efferent system. However, IC tuning shifts persist if the auditory cortex is inactivated after the conditioning phase, indicating that they do not depend on online corticofugal modulation (Gao and Suga 1998; Suga and Ma 2003). The importance of the auditory efferent system in mediating auditory plasticity has also been highlighted by studies showing that destruction of cortico-collicular neurons in ferrets selectively disrupts the ability to learn an auditory spatial localization task (Bajo et al. 2010; Bajo and King 2012).

Changes in the FFR after short-term training may occur either as a result of plastic changes that are local to brainstem circuits and/or as a result of online modulation by the auditory cortex via the efferent system. In the second case, FFR changes could not be considered the results of subcortical plasticity, as subcortical responses would passively reflect plasticity occurring at the cortical level. Krishnan and Gandour (2009) and Krishnan et al. (2010) have argued against this case, citing the fact that efferent activation is sluggish with respect to FFR responses that instead have short onset latencies (6–9 ms). For example, one well-studied efferent effect, the medial olivo-cochlear reflex, has an onset latency of about 25 ms (Backus and Guinan 2006). If the stimuli used for probing training effects had a static  $F_0$ , online efferent effects in this latency range could be useful to track the  $F_0$  of later portions of the stimuli. However, it is unlikely that efferent effects in this latency range would be of much help in tracking the time-varying  $F_0$  contours of the stimuli used in many FFR training studies unless the top-down efferent modulation used a memory trace of the stimulus to predict its time-varying  $F_0$  and facilitate responses phase locked to the predicted  $F_0$ . Chandrasekaran et al. (2014) have emphasized the idea that efferent

feedback may modulate brainstem activity on the basis of predictive coding schemes that attempt to anticipate the ongoing stimulation on the basis of preceding regularities in the stimulus stream. While there is disagreement on the relevance of efferent modulation for FFR changes after the training phase, most authors think that the efferent system is likely to be involved during the training phase. During the training phase efferent feedback could modulate brainstem activity online, as mentioned before, but the feedback could also be delayed and contingent on the outcome (correct or incorrect response) of a training trial.

The increased synchrony of FFR responses to stimulus-related periodicities induced by training and the apparent increase in SNR-based or autocorrelation-based measures may reflect either a greater accuracy of phase locking of single fibers to the periodicities in the stimulus or a greater proportion of fibers phase locking to stimulus-related periodicities. An increase in the proportion of neurons phase locking to stimulus-related periodicities may be the result of either the recruitment of additional fibers phase locking to these periodicities or the inhibition of fibers firing at different periods. Because the electroencephalogram reflects the summed response of many neurons, scalp-recorded FFRs cannot distinguish between these options. Nonetheless, it would be interesting to determine if increases in FFR SNR at the signal frequencies (or increases in ACF values at time lags related to the signal frequencies) after training are due to increased responses at the signal frequencies or decreased responses at the noise frequencies. It should be noted that a decreased response at the noise frequencies could reflect neuroplastic changes caused by training (e.g., inhibition of fibers not phase locked to the stimulus periodicities) but may also reflect a generic reduction of background physiological noise unrelated to neuroplastic changes, which could be caused instead by the participants being in a more relaxed state in the post-training FFR session. The use of a control group protects against this confound because participants in the control group should also be more relaxed at the post-training FFR session. However, this protection is not full because the greater familiarity that participants in the training group have with the stimuli used during the FFR recording may lead them to a more relaxed state compared to participants in the control group. Future studies should address this issue to ensure that measured FFR training effects truly reflect neuroplastic changes.

## 4.5 Summary

The research summarized in this chapter suggests that subcortical temporal processing changes as a result of training in pitch discrimination/identification tasks and in speech perception tasks. Training in these tasks leads to better tracking of stimulus-related periodicities and/or better timing of neural responses to sounds. This research, along with research on contextual effects on brainstem auditory processing (see Escera, Chap. 5) changes the view of auditory brainstem nuclei as static processing modules. Brainstem auditory processing appears instead to be dynamic because it shows short-term adaptation to the stimulus context, as well as

long-term training induced plasticity. In a way this is not surprising given the richness and complexity of subcortical auditory processing compared to the more circumscribed role that subcortical processing plays for other senses such as vision (King and Nelken 2009).

The idea that subcortical auditory processing can be modified by short-term training has important practical implications. Subcortical auditory processing is disrupted in certain populations of listeners such as children with language-based learning impairments and the elderly. Some of the studies reviewed in this chapter provide promising evidence that some of these deficits can be ameliorated by perceptual training. Further research is necessary to explore the full potential of auditory training to address nonperipheral hearing deficits and to find optimal training paradigms whose benefits translate to improvements in everyday communication settings that are maintained over time.

Although large inter-individual differences in the amount of perceptual learning are often noted in the literature, few studies have been designed to understand the causes of these differences. The results of Chandrasekaran et al. (2012) and Song et al. (2012), showing relationships between pre-training FFR metrics and perceptual learning, represent nice examples of how electrophysiological studies can shed light on the nature of these inter-individual differences. More studies are necessary to better understand the nature of inter-individual differences in perceptual learning. This would be important not only to improve theoretical understanding of perceptual learning but also for the development of individually tailored training programs in applied settings.

Much work remains to be done in order to understand how FFR plasticity occurs, both at the level of single neurons and at the level of neural circuits. This work will likely need to unravel complex interactions between subcortical nuclei and the efferent feedback from cortical areas. Other aspects of FFR plasticity that have not been fully explored are its degree of specificity with respect to particular stimulus features and its degree of generalization to untrained stimuli. Short-term training studies, although expensive in terms of time and resources, will be fundamental to answering these questions.

**Compliance with Ethics Requirements** Samuele Carcagno and Christopher J. Plack declared that they had no conflict of interest.

## References

- Amitay, S., Zhang, Y.-X., Jones, P. R., & Moore, D. R. (2014). Perceptual learning: Top to bottom. *Vision Research*, *99*, 69–77.
- Ananthanarayan, A. K., & Durrant, J. D. (1992). The frequency-following response and the onset response: Evaluation of frequency specificity using a forward-masking paradigm. *Ear and Hearing*, *13*(4), 228–232.
- Andersen, G. J., Ni, R., Bower, J. D., & Watanabe, T. (2010). Perceptual learning, aging, and improved visual performance in early stages of visual processing. *Journal of Vision*, *10*(13), article 4, 1–13.

- 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., Parbery-Clark, A., White-Schwoch, T., Dreihobl, S., & Kraus, N. (2013a). Effects of hearing loss on the subcortical representation of speech cues. *The Journal of the Acoustical Society of America*, 133(5), 3030–3038.
- Anderson, S., White-Schwoch, T., Choi, H. J., & Kraus, N. (2013b). Training changes processing of speech cues in older adults with hearing loss. *Frontiers in Systems Neuroscience*, 7. Doi:10.3389/fnsys.2013.00097
- Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013c). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences of the USA*, 110(11), 4357–4362.
- Anderson, S., White-Schwoch, T., Choi, H. J., & Kraus, N. (2014). Partial maintenance of auditory-based cognitive training benefits in older adults. *Neuropsychologia*, 62, 286–296.
- Atienza, M., Cantero, J. L., & Dominguez-Marín, E. (2002). The time course of neural changes underlying auditory perceptual learning. *Learning & Memory*, 9(3), 138–150.
- Backus, B. C., & Guinan, J. J. (2006). Time-course of the human medial olivocochlear reflex. *The Journal of the Acoustical Society of America*, 119(5), 2889–2904.
- Bajo, V. M., & King, A. J. (2012). Cortical modulation of auditory processing in the midbrain. *Frontiers in Neural Circuits*, 6. Doi:10.3389/fncir.2012.00114
- 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.
- Bharadwaj, H. M., & Shinn-Cunningham, B. G. (2014). Rapid acquisition of auditory subcortical steady state responses using multichannel recordings. *Clinical Neurophysiology*, 125(9), 1878–1888.
- Bidelman, G. M. (2015). Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. *Hearing Research*, 323, 68–80.
- Bones, O., & Plack, C. J. (2015). Losing the music: Aging affects the perception and subcortical neural representation of musical harmony. *The Journal of Neuroscience*, 35(9), 4071–4080.
- Bosnyak, D. J., Eaton, R. A., & Roberts, L. E. (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebral Cortex*, 14(10), 1088–1099.
- Bower, J. D., Watanabe, T., & Andersen, G. J. (2013). Perceptual learning and aging: Improved performance for low-contrast motion discrimination. *Frontiers in Psychology*, 4. Doi:10.3389/fpsyg.2013.00066
- Brown, M., Irvine, D. R. F., & Park, V. N. (2004). Perceptual learning on an auditory frequency discrimination task by cats: Association with changes in primary auditory cortex. *Cerebral Cortex*, 14(9), 952–965.
- Carcagno, S., & Plack, C. J. (2011a). Pitch discrimination learning: Specificity for pitch and harmonic resolvability, and electrophysiological correlates. *Journal of the Association for Research in Otolaryngology*, 12(4), 503–517.
- Carcagno, S., & Plack, C. J. (2011b). Subcortical plasticity following perceptual learning in a pitch discrimination task. *Journal of the Association for Research in Otolaryngology*, 12, 89–100.
- Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences*, 8(10), 465–471.
- Chandrasekaran, B., & Kraus, N. (2012). Biological factors contributing to reading ability: Subcortical auditory function. In A. A. Benasich & R. H. Fitch (Eds.), *Developmental dyslexia: Early precursors, neurobehavioral markers and biological substrates* (pp. 83–98). Baltimore: Paul H. Brookes Publishing.
- Chandrasekaran, B., Kraus, N., & Wong, P. C. M. (2012). Human inferior colliculus activity relates to individual differences in spoken language learning. *Journal of Neurophysiology*, 107(5), 1325–1336.
- Chandrasekaran, B., Skoe, E., & Kraus, N. (2014). An integrative model of subcortical auditory plasticity. *Brain Topography*, 27(4), 539–552.

- Clinard, C. G., & Tremblay, K. L. (2013). Aging degrades the neural encoding of simple and complex sounds in the human brainstem. *Journal of the American Academy of Audiology*, 24(7), 590–599.
- Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Hearing Research*, 264(1–2), 48–55.
- Dau, T. (2003). The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. *The Journal of the Acoustical Society of America*, 113(2), 936–950.
- de Cheveigné, A. (2005). Pitch perception models. In R. Fay, A. N. Popper, C. J. Plack, & A. J. Oxenham (Eds.), *Pitch: Neural coding and perception* (pp. 169–233). New York: Springer.
- Dediu, D., & Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and microcephalin. *Proceedings of the National Academy of Sciences of the U S A*, 104(26), 10944–10949.
- Deveau, J., & Seitz, A. R. (2014). Applying perceptual learning to achieve practical changes in vision. *Frontiers in Psychology*, 5(1166), 1–6.
- de Villiers-Sidani, E., Alzghoul, L., Zhou, X., Simpson, K. L., et al. (2010). Recovery of functional and structural age-related changes in the rat primary auditory cortex with operant training. *Proceedings of the National Academy of Sciences of the U S A*, 107(31), 13900–13905.
- Doshier, B. A., & Lu, Z. L. (1999). Mechanisms of perceptual learning. *Vision Research*, 39(19), 3197–3221.
- Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., & Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science*, 291(5510), 1969–1972.
- Drennan, W. R., Won, J. H., Dasika, V. K., & Rubinstein, J. T. (2007). Effects of temporal fine structure on the lateralization of speech and on speech understanding in noise. *Journal of the Association for Research in Otolaryngology*, 8(3), 373–383.
- Engen, K. J. V. (2012). Speech-in-speech recognition: A training study. *Language and Cognitive Processes*, 27(7–8), 1089–1107.
- Fu, Q.-J., & Galvin, J. J., 3rd. (2007). Perceptual learning and auditory training in cochlear implant recipients. *Trends in Amplification*, 11(3), 193–205.
- Gao, E., & Suga, N. (1998). Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proceedings of the National Academy of Sciences of the U S A*, 95(21), 12663–12670.
- Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, 31(5), 681–697.
- Gockel, H. E., Carlyon, R. P., Mehta, A., & Plack, C. J. (2011). The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch. *Journal of the Association for Research in Otolaryngology*, 12(6), 767–782.
- Gockel, H. E., Krugliak, A., Plack, C. J., & Carlyon, R. P. (2015). Specificity of the human frequency following response for carrier and modulation frequency assessed using adaptation. *Journal of the Association for Research in Otolaryngology*, 16(6), 747–762.
- Goldstone, D. R. L., Braithwaite, D. W., & Byrge, L. A. (2012). Perceptual learning. In P. D. N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 2580–2583). New York: Springer.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1), 14–23.
- Grothe, B., Pecka, M., & McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiological Reviews*, 90(3), 983–1012.
- Hensch, T. K. (2004). Critical period regulation. *Annual Review of Neuroscience*, 27, 549–579.
- Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron*, 76(3), 486–502.
- Hornickel, J., Zecker, S. G., Bradlow, A. R., & Kraus, N. (2012). Assistive listening devices drive neuroplasticity in children with dyslexia. *Proceedings of the National Academy of Sciences of the U S A*, 109(41), 16731–16736.



- Jones, P. R., Moore, D. R., Shub, D. E., & Amitay, S. (2015). The role of response bias in perceptual learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1456–1470.
- King, A. J., & Nelken, I. (2009). Unraveling the principles of auditory cortical processing: Can we learn from the visual system? *Nature Neuroscience*, *12*(6), 698–701.
- Kraus, N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Annals of the New York Academy of Sciences*, *1169*, 543–557.
- 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.
- Krishnan, A. (2007). Frequency-following response. In R. F. Burkard, J. J. Eggermont, & M. Don (Eds.), *Auditory evoked potentials: Basic principles and clinical applications* (pp. 313–333). Philadelphia: Lippincott Williams.
- Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*, *110*(3), 135–148.
- Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, *25*(1), 161–168.
- Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, *23*(1), 81–95.
- Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America*, *96*(4), 2076–2087.
- Marmel, F., Linley, D., Carlyon, R. P., Gockel, H. E., et al. (2013). Subcortical neural synchrony and absolute thresholds predict frequency discrimination independently. *Journal of the Association for Research in Otolaryngology*, *14*(5), 757–766.
- Menning, H., Roberts, L. E., & Pantev, C. (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *NeuroReport*, *11*(4), 817–822.
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*(1–2), 36–47.
- Mollon, J. D., & Danilova, M. V. (1996). Three remarks on perceptual learning. *Spatial Vision*, *10*(1), 51–58.
- Monaghan, P., Metcalfe, N. B., & Ruxton, G. D. (1998). Does practice shape the brain? *Nature*, *394*(6692), 434.
- Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology*, *9*(4), 399–406.
- Moore, D. R., & Shannon, R. V. (2009). Beyond cochlear implants: Awakening the deafened brain. *Nature Neuroscience*, *12*(6), 686–691.
- Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: A problem of significance. *Nature Neuroscience*, *14*(9), 1105–1107.
- Petrov, A. A., Doshier, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review*, *112*(4), 715–743.
- Plack, C. J., & Oxenham, A. J. (2005). The Psychophysics of Pitch. In R. Fay, A. N. Popper, C. J. Plack, & A. J. Oxenham (Eds.), *Pitch: Neural coding and perception*. New York: Springer.
- Polat, U., Schor, C., Tong, J.-L., Zomet, A., et al. (2012). Training the brain to overcome the effect of aging on the human eye. *Scientific Reports*, *2*, 278. Doi:10.1038/srep00278
- 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. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 336(1278), 367–373.
- Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, 156(1), 95–103.
- Sheehan, K. A., McArthur, G. M., & Bishop, D. V. M. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Cognitive Brain Research*, 25(2), 547–553.
- Skoe, E., Chandrasekaran, B., Spitzer, E. R., Wong, P. C. M., & Kraus, N. (2014). Human brainstem plasticity: The interaction of stimulus probability and auditory learning. *Neurobiology of Learning and Memory*, 109, 82–93.
- Skuk, V. G., & Schweinberger, S. R. (2014). Influences of fundamental frequency, formant frequencies, aperiodicity, and spectrum level on the perception of voice gender. *Journal of Speech, Language, and Hearing Research*, 57(1), 285–296.
- Song, J. H., Nicol, T., & Kraus, N. (2011a). Test-retest reliability of the speech-evoked auditory brainstem response. *Clinical Neurophysiology*, 122(2), 346–355.
- Song, J. H., Nicol, T., & Kraus, N. (2011b). Reply to Test-retest reliability of the speech-evoked ABR is supported by tests of covariance. *Clinical Neurophysiology*, 122(9), 1893–1895.
- Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: Biological mechanisms. *Cerebral Cortex*, 22(5), 1180–1190.
- Song, J. H., Skoe, E., Wong, P. C. M., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, 20(10), 1892–1902.
- Sowden, P. T., Davies, I. R., & Roling, P. (2000). Perceptual learning of the detection of features in X-ray images: A functional role for improvements in adults' visual sensitivity? *Journal of Experimental Psychology: Human Perception and Performance*, 26(1), 379–390.
- Suga, N., & Ma, X. (2003). Multiparametric corticofugal modulation and plasticity in the auditory system. *Nature Reviews Neuroscience*, 4(10), 783–794.
- Tong, Y., Melara, R. D., & Rao, A. (2009). P2 enhancement from auditory discrimination training is associated with improved reaction times. *Brain Research*, 1297, 80–88.
- Tremblay, K. L., Ross, B., Inoue, K., McClannahan, K., & Collet, G. (2014). Is the auditory evoked P2 response a biomarker of learning? *Frontiers in Systems Neuroscience*, 8(28), 1–13.
- Tsodyks, M., & Gilbert, C. (2004). Neural networks and perceptual learning. *Nature*, 431(7010), 775–781.
- Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: Mechanisms of plasticity in the auditory brainstem. *Neuron*, 62(4), 463–469.
- Watanabe, T., & Sasaki, Y. (2015). Perceptual learning: Toward a comprehensive theory. *Annual Review of Psychology*, 66, 197–221.
- Weinberger, N. M. (2008). Cortical plasticity in associative learning and memory. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference*. Vol. 3. *Memory systems* (pp. 187–218). Oxford: Academic Press.
- 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(4), 420–422.
- Wright, B. A., & Zhang, Y. (2009). A review of the generalization of auditory learning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1515), 301–311.
- Wright, B. A., Wilson, R. M., & Sabin, A. T. (2010). Generalization lags behind learning on an auditory perceptual task. *The Journal of Neuroscience*, 30(35), 11635–11639.
- Xu, Y., Krishnan, A., & Gandour, J. T. (2006). Specificity of experience-dependent pitch representation in the brainstem. *NeuroReport*, 17(15), 1601–1605.