JARO

Journal of the Association for Research in Otolaryngology

Subcortical Plasticity Following Perceptual Learning in a Pitch Discrimination Task

Samuele Carcagno[1,2] and Christopher J. Plack[3]

[1]Department of Psychology, Lancaster University, Lancaster, LA1 4YF, UK
[2]CNRS and Université de Bordeaux, BP 63, 146 Rue Leo Saignat, F-33076, Bordeaux, France
[3]Division of Human Communication and Deafness, The University of Manchester, Manchester, M13 9PL, UK

ABSTRACT

Practice can lead to dramatic improvements in the discrimination of auditory stimuli. In this study, we investigated changes of the frequency-following response (FFR), a subcortical component of the auditory evoked potentials, after a period of pitch discrimination training. Twenty-seven adult listeners were trained for 10 h on a pitch discrimination task using one of three different complex tone stimuli. One had a static pitch contour, one had a rising pitch contour, and one had a falling pitch contour. Behavioral measures of pitch discrimination and FFRs for all the stimuli were measured before and after the training phase for these participants, as well as for an untrained control group ($n=12$). Trained participants showed significant improvements in pitch discrimination compared to the control group for all three trained stimuli. These improvements were partly specific for stimuli with the same pitch modulation (dynamic vs. static) and with the same pitch trajectory (rising vs. falling) as the trained stimulus. Also, the robustness of FFR neural phase locking to the sound envelope increased significantly more in trained participants compared to the control group for the static and rising contour, but not for the falling contour. Changes in FFR strength were partly specific for stimuli with the same pitch modulation (dynamic vs. static) of the trained stimulus. Changes in FFR strength, however, were not specific for stimuli with the same pitch trajectory (rising vs. falling) as the

trained stimulus. These findings indicate that even relatively low-level processes in the mature auditory system are subject to experience-related change.

Keywords: auditory training, FFR, F0 discrimination, evoked potentials

INTRODUCTION

Sounds whose waveforms repeat periodically elicit a sensation of pitch, which plays a fundamental role in the perception of speech and music, as well as in the segregation of concurrent sound sources (Plack and Oxenham 2005a). Auditory nerve fibers "phase lock" to the waveform of pure tones, that is, neural firing tends to occur at the same time during each cycle of the sound wave. Several models of pitch assume that this temporal information, which is preserved in upper brainstem structures for fundamental frequencies (F0s) up to several hundred hertz (Liu et al. 2006), is used to encode pitch (e.g., Meddis and O'Mard 2006; McLachlan 2009).

Musicians (Wong et al. 2007; Bidelman et al. 2009) and speakers of a tone language (Krishnan et al. 2005, 2009a, b; Swaminathan et al. 2008) show more robust subcortical "phase locking" in response to pitch-evoking sounds compared to English speakers without musical experience. This has been demonstrated using scalp recordings of the frequency following response (FFR), an evoked potential which reflects neural phase locking of brainstem nuclei (inferior colliculus and lateral lemniscus, Smith et al. 1975; Gardi et al. 1979) to the envelope of a sound. The

Correspondence to: Samuele Carcagno · CNRS and Université de Bordeaux, BP 63 · 146 Rue Leo Saignat, F-33076, Bordeaux, France. Telephone: +33-5-57571651; email: samuele.carcagno@u-bordeaux2.fr

enhancement of the FFR in musicians and speakers of a tone language has been explained as a result of subcortical plasticity driven by the extensive practice these populations of listeners have in identifying and discriminating sounds on the basis of pitch. However, FFR differences between expert and naive pitch listeners may be caused by factors other than neural plasticity (Monaghan et al. 1998), such as genetic predispositions. Moreover, assuming that FFR enhancements in expert pitch listeners reflect subcortical plasticity, it remains unclear whether this plasticity is limited to a critical developmental period, or is retained also in adulthood. A more direct approach to the study of neural plasticity consists in measuring neural activity before and after a period of training. The only study using this approach in adults found more robust FFR phase locking to the waveform of one out of three tones in a group of English speakers following a period of training in an auditory identification task with pseudo-words with the same pitch contour as Mandarin tones (Song et al. 2008). The lack of a control group in this study, however, does not allow the unequivocal conclusion that FFR changes between the pre- and post-training recordings were due to auditory training per se. One purpose of the present study was to provide a more rigorous test of the experience-dependent plasticity of the FFR. To this end, we compared FFR changes between a group of adult listeners following an auditory training protocol of ten 1-h sessions and a control group that did not receive any training. FFR enhancements in Mandarin speakers have been found to be specific for stimuli with the same pitch contour as Mandarin tones (Xu et al. 2006; Krishnan et al. 2009a). To assess the specificity of FFR training effects with respect to pitch contour, we trained participants with one of three stimuli respectively with a rising, falling, or static pitch contour, and assessed FFR changes after training for all three stimuli.

Changes in performance in the behavioral discrimination task were also assessed for all stimuli.

METHODS

Participants

Thirty-nine participants (16 females, 35 right handed, two ambidextrous) completed the experiment. The participants ranged in age between 19 and 35 years (mean=23, SD=2). They all had normal hearing for both ears with absolute pure tone thresholds below 20 dB HL at octave frequencies from 250 to 8000 Hz. None had prior experience in psychoacoustic tasks or musical training. All participants gave written informed consent and were paid an hourly wage for their participation in the experiment. All procedures of the study were approved by the Department of Psychology Ethics Committee, Lancaster University.

Stimuli

In order to assess the specificity of training effects with respect to the shape of the pitch contour, the experimental group was subdivided into three groups (G-Up, G-Down, G-Static) trained respectively with harmonic complex tones with a rising (S-Up), falling (S-Down), or static (S-Static) pitch contour (see Fig. 1). In order to maximize the chances of observing possible training effects on the FFR, the average F0 of the stimuli, and the F0 excursion of the dynamic pitch contour stimuli were similar to those of the Mandarin tones which have been used in previous research on FFR plasticity (e.g., Krishnan et al. 2005).

The tasks used for the training protocol were the same as the tasks used for assessing behavioral performance pre- and post-testing. For the static tone, participants were required to discriminate between a standard stimulus with a fixed 140 Hz F0 and a comparison stimulus with a lower F0 (see Fig. 2A). The duration of the stimulus was 450 ms, including 10 ms rise and fall cosine ramps. The dynamic tones were derived by modulating the F0 of the 140 Hz static tone with half a cycle of a sine wave (80% modulation depth). The task for these tones required the discrimination between a standard stimulus with a 400 ms half-cycle FM and a comparison stimulus with a shorter half-cycle FM (see Fig. 2B). The starting phase of the modulation was 1.5π radians for the S-Up tone, and 0.5π radians for the S-Down tone. The modulation was temporally centered in the middle of the stimuli. In order to encourage listeners to attend to changes in F0 rate, rather than to F0 differences at specific times of the stimuli, the duration of the unmodulated segments at the beginning and at the end of the stimuli, was extended by a value drawn
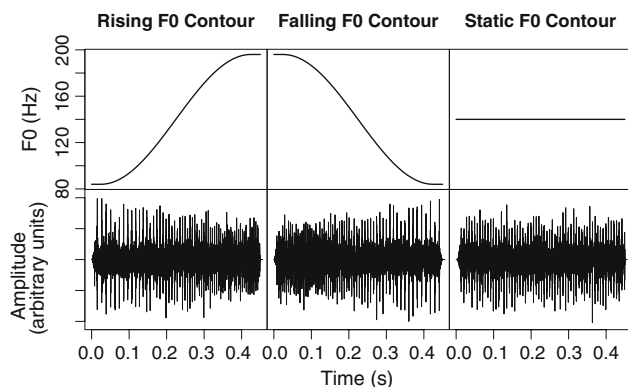
FIG. 1. F0 contours (*top row*) and waveforms (*bottom row*) of the three complex tones used in the experiment.
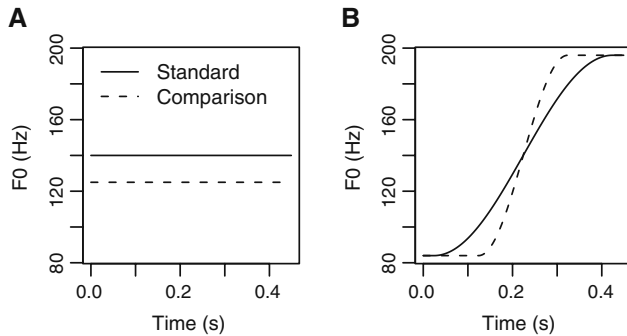
A

B

FIG. 2. Illustration of the psychophysical tasks. **A** For the static stimulus, participants were required to discriminate between a standard stimulus with a fixed 140 Hz F0 and a comparison stimulus with a lower F0. **B** F0 contour of the rising stimulus (S-Up). For the dynamic stimuli, participants were required to discriminate between a standard stimulus with a 400 ms half-cycle FM and a comparison stimulus with a shorter half-cycle FM. The start and end F0 points were the same for the standard and comparison stimuli.

randomly from a uniform distribution between 0 and 50 ms, only for the behavioral task. The harmonics of all the tones were added in sine phase. All tones were filtered between 2 and 3 kHz, and contained only unresolved harmonics, to ensure that listeners were using a temporal mechanism to extract the missing fundamental of the tones (Plack and Oxenham 2005b). This had the purpose of forcing the listeners to use the temporal cues which are thought to be encoded by the FFR, rather than place cues, to perform the task. The tones had an overall level of 80 dB SPL. In order to keep the overall level of the tones constant, the amplitude of the F0-modulated tones was modulated by a factor of $\sqrt{F0_t/F0_{\text{base}}}$, where $F0_t$ is the F0 of the complex at time t and $F0_{base}$ the F0 of the complex at a zero crossing of the FM. A low-pass noise with a 1.9 kHz cutoff was added to all the tones in order to mask possible combination tones, its spectrum level was 5 dB below the spectrum level of the tones. Another band of noise with the same spectrum level was added in a higher frequency region between 3.1 and 4.2 kHz to limit the contribution of cochlear regions with characteristic frequencies above the frequencies of the tones to the FFR response (Ananthanarayan and Durrant 1992; Dau 2003). The noise side bands were generated independently for each presentation during the behavioral sessions, but corresponded to frozen samples for the electrophysiological sessions.

For the behavioral sessions, the stimuli were generated digitally with 32-bit resolution and a 48-kHz sampling rate on a Macintosh workstation. The stimuli were played through an M-AUDIO Firewire 410 DAC and presented binaurally via Sennheiser HD580 headphones. For the FFR sessions, all stimuli were generated digitally with 16-bit resolution and a

40-kHz sampling rate. The stimulus files were played through a DAC included in the evoked potentials data acquisition system (Intelligent Hearing Systems–Smart EP), and presented binaurally through mu-metal shielded ER2 insert earphones. Binaural FFRs have greater amplitude than FFRs recorded monaurally (Clark et al. 1997). The signal-to-noise ratio for binaural FFRs should therefore be greater than for monaural FFRs, and allow for more accurate FFR measurement.

Design

Participants were randomly assigned to one of three experimental groups (G-Up, $n=9$; G-Down, $n=9$; G-Static, $n=9$) or to a control group (G-Control, $n=12$). All participants took part in a preliminary session during which they were familiarized with the stimuli and procedures of the experiment by running two blocks of the discrimination task for each stimulus. An additional familiarization block for each stimulus was run during the first behavioral discrimination thresholds session. Pre-training FFRs and behavioral discrimination thresholds were then measured in separate successive sessions. During the training phase, participants of the experimental groups ran ten sessions of the auditory discrimination task. During each training session, participants of the G-Up and G-Down groups completed 20 blocks of the discrimination task with the S-Up and S-Down stimuli, respectively, while participants of the G-Static group completed 18 blocks of the discrimination task with the S-Static stimulus. The mean duration of the training phase was 27 days. Participants of the control group waited for a similar amount of time (mean 32 days) without receiving any training. After the training phase, FFRs and behavioral thresholds were measured again in three separate sessions.

Behavioral procedure

Pre- and post-testing thresholds for the discrimination task were measured with a three-interval, three-alternative forced-choice task, using an adaptive procedure. On each trial, three observation intervals separated by 500 ms, each containing a harmonic complex tone, were presented. For the static stimulus, two observation intervals (standard intervals), chosen randomly, were assigned a complex with a fixed F0 of 140 Hz, the other observation interval (comparison interval) contained a complex tone of a lower F0 which was varied adaptively. For the dynamic stimuli, two observation intervals (standard intervals), chosen randomly, were assigned a complex with a fixed FM duration of 400 ms, the other observation interval (comparison interval) contained a complex tone with a shorter FM duration, which was varied adaptively.

The listener was asked to indicate by a key press on a numeric keypad which of the tones sounded different from the other two (odd-one-out task). Feedback was always provided at the end of each trial through the presentation of a colored light on the computer screen. For the static stimulus, a two-down one-up adaptive rule tracking the 70.7% correct point on the psychometric function was used (Levitt 1971). The percentage F0 difference between the complex tones in the standard and comparison observation intervals was initially set at 20%, and was increased (after an incorrect response) and decreased (after two consecutive correct responses) by a factor of 2 for the first four turnpoints and by a factor 1.414 thereafter. The maximum percentage F0 difference allowed was 80%. Sixteen turnpoints were measured for each block of trials and the threshold estimate was taken as the geometric mean of the last 12. For the dynamic stimuli a modified two-down one-up adaptive rule was used. The percentage FM duration difference between the complex tones in the standard and comparison observation intervals was initially set at 50%, and was increased after an incorrect response, and decreased after two consecutive correct responses, by a factor of 2 for the first four threshold estimation points and by a factor of 1.414 thereafter. The threshold estimation points were either turnpoints, or points at which the listener had given an incorrect response after reaching the FM duration difference limit (99%). A block was terminated after 16 threshold estimation points were collected. The threshold estimate was taken as the geometric mean of the last 12 of such points. For both static and dynamic stimuli, auditory discrimination thresholds for each stimulus were computed as the geometric mean of the threshold estimates measured in five blocks of trials. The change in performance across the pre- and post-thresholds assessment sessions was quantified as the ratio of the pre- to post-testing threshold.

FFR recordings

Participants reclined comfortably in a double-walled sound-attenuating booth. They were instructed to relax and refrain from extraneous body movements. The electroencephalogram (EEG) was recorded differentially between gold-plated scalp electrodes placed on the midline of the forehead at the hairline and the seventh cervical vertebra. Another electrode placed on the mid-forehead served as the common ground. The interelectrode impedances were maintained below 1 kΩ. The EEG signal was recorded with an 8 kHz sampling rate, bandpass filtered from 50 to 3,000 Hz, and amplified by a factor of 150,000. The stimuli were played with a repetition rate of two per second and were presented in blocks of 256, in alternating polarity (half in rarefaction and half in condensation polarity). The sum of the waveforms recorded in opposite polarities was used for the analyses. Epochs with voltage changes exceeding 29 μV were automatically discarded and the trial repeated. Seven blocks were repeated for each stimulus. The order of the blocks was randomized. The online EEG activity was monitored and if the EEG was noisy during a certain block, it was noted and discarded from subsequent analyses. Either the last six blocks, or the six blocks remaining after discarding blocks with noisy recordings, were used for the analyses. The overall duration of a session, including electrode placement, was about 1 h and 30 min. The FFR waveforms were bandpass filtered offline with digital finite impulse response filters between 50 and 1,900 Hz. The high-frequency cutoff was chosen in order to ensure that harmonic components generated by the transducers were not contaminating the recording

Statistical analyses

The changes in behavioral thresholds (ratio of the thresholds measured at the pre- and post-threshold assessment session) were log transformed to improve the normality of the data. Before computing test statistics, means and standard deviations of each dependent measure (change in behavioral thresholds, FFR strength), were computed for each combination of stimulus (S-Up, S-Down, S-Static) per group (G-Up, G-Down, G-Static, G-Control). Data points falling beyond ±2 standard deviations of the group mean for a given stimulus were considered outliers. All the data of a participant with one or more outliers in a given dependent measure were discarded from the analyses of that dependent measure only. Outliers were present only in the FFR measure for four participants (one in the G-Up group, two in the G-Static group, and one in the G-Control group). All comparisons were planned, except where explicitly stated, and the reported p values are uncorrected. When the test statistic involved a t test between independent samples, the Fligner–Killeen test of the homogeneity of variances between the two groups (Conover et al. 1981) was first performed. In the case of unequal variances between the two groups the Welch–Satterthwaite approximation to the degrees of freedom (Satterthwaite 1946) was applied. Since the expected direction of change for the dependent measures was known, all the t tests were run as one-tailed tests, except where explicitly stated. For the analysis of the correlations between the behavioral and physiological measures, we employed a non-parametric procedure (Spearman's rank correlation)

that does not rely on the assumption of normally distributed data. Since the expected sign of the correlations was known, their significance was tested with one-tailed t tests.

RESULTS

Learning curves

The FM duration difference limen (DL) and F0 DL across all training sessions are displayed in Figure 3. Before looking at the threshold changes across the pre- and post-testing sessions, it is worth examining the performance of the participants during the first behavioral session, as well as the performance improvements across training sessions. The F0 DLs measured in the first session for the static tone were higher than the thresholds measured in naive listeners for similar stimuli in another study (Grimault et al. 2002). This is at least in part due to the fact that the three-interval, three-alternative forced-choice procedure used in the present experiment tracked a higher level of performance (a higher d′ value) than the three-interval, two-alternative forced-choice procedure used in the earlier study. Performance on the F0 discrimination task in the present experiment may also have been reduced due to the presence of an

additional band of noise in a high frequency region, added to limit the contribution of cochlear regions with characteristic frequencies above the frequencies of the tones to the FFR response.

FM duration DLs during the first behavioral session for the stimuli with a dynamic F0 contour appear to be very high, close to the maximum FM duration difference allowed (99%). Since listeners were often hitting the maximum FM duration difference, and also because of the modified adaptive procedure used to estimate thresholds in the FM duration discrimination task, it is not known to which point on the psychometric function the threshold estimate corresponds. It is possible that some participants were performing at chance level on the FM duration discrimination task. To test whether or not this was the case, a simulation of a listener performing at chance level for all the levels of the adaptive difference on the FM duration discrimination task was run. Figure 4 shows the probability distribution of the FM duration DL estimate for a five blocks session, for such a listener. The distribution was computed by running the simulation for 10,000 sessions comprising five blocks of the FM duration discrimination task each. The results of the simulation indicate that the probability that a listener performing at chance level would achieve a FM duration DL below 90% is 9.3%.
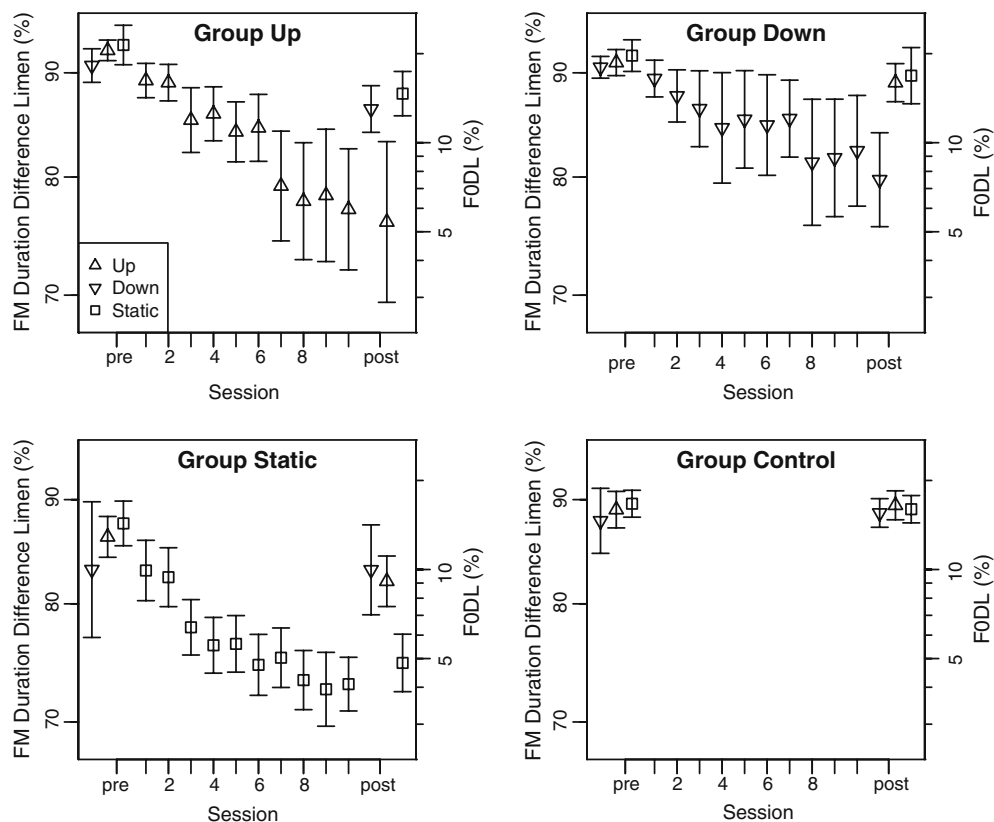

FIG. 3. Mean F0 and FM duration DLs across sessions ±1 standard error of the mean. The *triangles pointing up and down* plot the DLs for the S-Up and S-Down stimuli respectively. The *square* plots the DLs for the S-Static tone, and its scale should be referred to the axis on the *right*.

Out of the total 78 threshold estimates (39 listeners×2 stimuli), 33 were below 90%. Thus, it is likely that many listeners performed above chance level in the FM duration discrimination task at the first behavioral session, although it is possible that a number of listeners were performing at chance level.

The DL changes over the training sessions displayed in Figure 3 suggest that both the groups trained on the FM duration discrimination task and the group trained on the F0 discrimination task showed a protracted decrease in thresholds across the training sessions. For the groups trained on the FM duration discrimination task this was confirmed by a repeated-measures analysis of variance (ANOVA) on the log-transformed FM duration DLs, with SESSION (1–10) as within and GROUP (G-Up, G-Down) as between-subjects factors. This analysis revealed a significant main effect of SESSION [$F(9,144)$=4.363, p<0.001], while the main effect of GROUP [$F(1,16)$=0.138, p=0.715] and the GROUP x SESSION interaction [$F(9,144)$=0.663, p=0.741] were not significant. These results indicate that thresholds decreased over the training sessions for both the G-Up and G-Down groups. The results of a univariate repeated-measures ANOVA on the log-transformed F0 DLs for the group trained on the F0 discrimination task revealed a significant effect of SESSION [$F(9,72)$=10.363, p<0.001] as well.

Threshold changes between the pre- and post-testing sessions

The ratio of the pre- to post-testing threshold for all stimuli and groups is displayed in Figure 5A, B showing the ratio of the pre- to post-testing threshold only for a set of selected planned contrasts. These
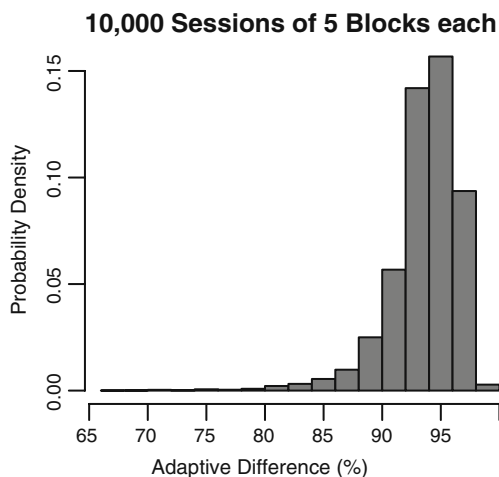
10,000 Sessions of 5 Blocks each

FIG. 4. Probability distribution of FM duration DLs in the FM duration discrimination task for a listener performing at chance at all levels of the FM duration difference between the stimulus in the standard and comparison intervals.

contrasts compare changes for each trained group on the trained stimulus, to changes in the control group for the same stimulus. Each of the experimental groups showed a significantly greater performance improvement on the trained stimulus (post vs. pre) compared to the improvement for the same stimulus in the control group [G-Up, $t(19)$=2.498, p<0.011; G-Down, $t(13.06)$=2.360, p<0.017; G-Static, $t(19)$=4.841, p<0.001]. This confirms that the auditory training protocol was effective. A significant interaction [$F(1,16)$=8.527, p<0.010] in a two-way mixed ANOVA, with GROUP (G-Up, G-Down) as the between-subject factor, and STIMULUS (S-Up, S-Down) as the within-subjects factor, indicated that learning in the discrimination task with the dynamic pitch stimuli was partly specific to the trajectory of the pitch contour (rising vs. falling). Collapsing the dynamic pitch stimuli (S-Up and S-Down) into a single STIMULUS level (S-Dynamic) and the dynamic groups (G-Up and G-Down) into a single GROUP level (G-Dynamic), showed that learning was also partly specific to the modulation of the pitch contour (dynamic vs. static), as indicated by a significant interaction [$F(1,25)$=19.982, p<0.001] in a two-way mixed ANOVA with GROUP (G-Dynamic, G-Static) as the between-subjects factor, and STIMULUS (S-Dynamic, S-Static) as the within-subjects factor.

Plasticity at the subcortical Level

A 100-ms segment of an FFR trace from a representative participant is shown in Figure 6A, and its corresponding autocorrelation function (ACF) is shown in Figure 6B. The ACF for each time lag τ=l/f where l is the lag expressed as number of samples, and f is the sampling frequency, was computed as $ACF(\tau) = \sum_{i=1}^{N} s(i)s(i-l)$, where i is the sample number, N is the total number of points in the signal and $s(x)$ is the amplitude of the FFR waveform at sample number x. The autocorrelation values were then normalized by dividing them by the maximum autocorrelation value. The FFR trace has a clear periodic component corresponding to the period of the stimulus, which results in a peak at the time lag corresponding to this period in the ACF. This indicates phase locking to the F0 of the stimulus. In order to quantify the strength of phase locking to the F0, we analyzed the FFR responses with an algorithm that extracted the value of the normalized ACF at the time lag corresponding to the inverse of the average F0 ($ACF_{1/F0}$) of the stimuli in 64 ms non-overlapping time frames. An autocorrelogram is a plot of the value of the ACF as a function of both period and time. Difference grand average autocorrelograms between the pre- and post-FFR assessment sessions, showing
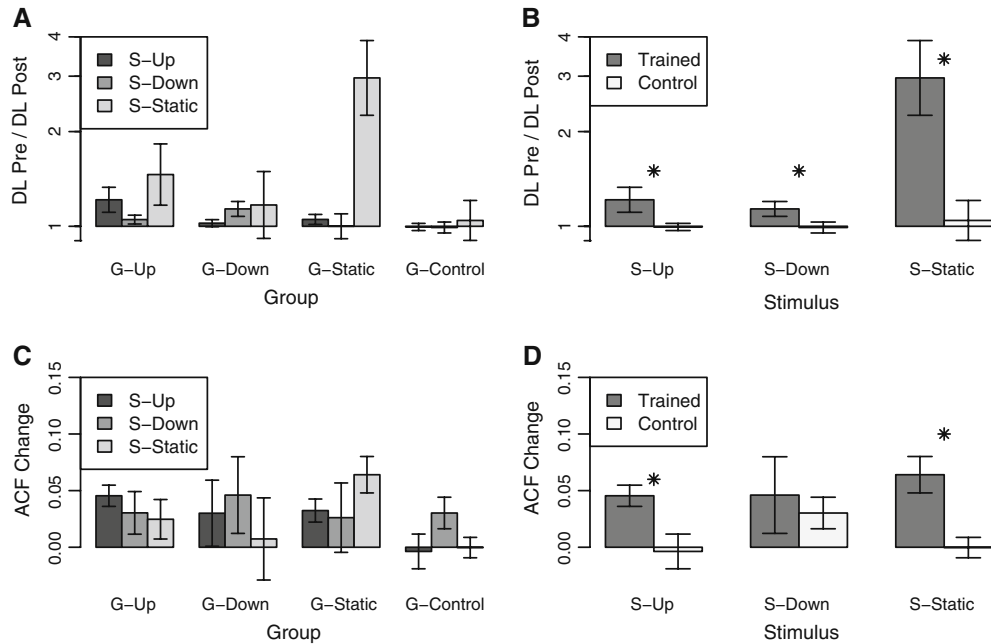
FIG. 5. Changes in the behavioral and FFR measures. **A** Ratio of the thresholds measured at the pre- and post-behavioral sessions. Values above one indicate improvements. **B** Same as **A**, but for a set of planned contrasts comparing changes for each trained group on the trained stimulus, to changes in the control group for the same stimulus. **C** Difference of the ACF value at the time lags corresponding to the inverse of the F0 of the stimuli between post- and pre-FFR assessment sessions. **D** Same as **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.
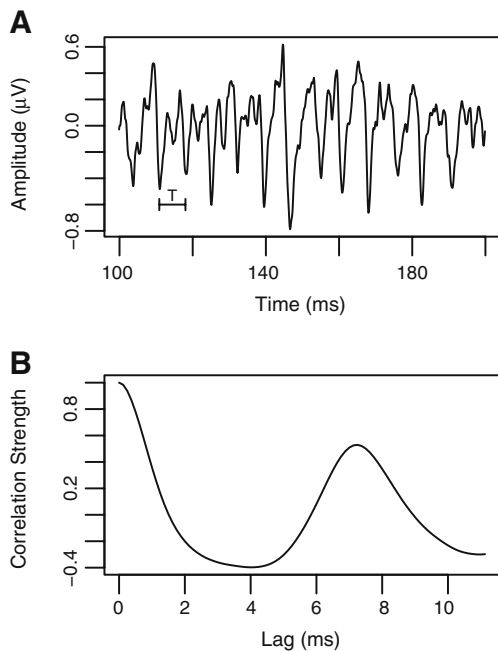

FIG. 6. **A** Segment of an FFR trace (average of 1,536 repetitions) recorded from a representative participant in response to the static F0 stimulus during the pre-testing session. The period $T=7.1$ ms of the static F0 stimulus is indicated by the *bar at the bottom left*. As a result of phase locking, the FFR trace tends to repeat with the same period. **B** Normalized ACF function of the FFR trace shown in **A**, the periodicity of the FFR trace generates a peak in the ACF function at the time lag corresponding to the period.

differences in ACF value between the two sessions are displayed in Figure 7. Figure 8 shows the $ACF_{1/F0}$ values for all stimuli and groups at the pre- and post-testing sessions, separately. For all the experimental groups, the value of the $ACF_{1/F0}$ tended to increase after training for the trained stimuli. Figure 5C shows the difference in the $ACF_{1/F0}$ value between the pre- and post-testing sessions for all stimuli and groups. This same difference is shown only for a set of selected planned contrasts in Figure 5D. The $ACF_{1/F0}$ increased significantly more for the trained stimuli than for the same stimuli in the control group for the G-Up [$t(17)=2.485$, $p<0.012$], and G-Static [$t(16)=3.782$, $p<0.001$] groups, while the effect failed to reach significance for the G-Down group [$t(18)=0.464$, $p=0.324$]. For the stimuli with a dynamic pitch contour, the specificity of the FFR training effects with respect to pitch trajectory (rising vs. falling) did not reach significance [$F(1,15)=0.988$, $p=0.336$]. FFR training effects were however specific to the modulation (dynamic vs. static) of the pitch contour [$F(1,22)=6.323$, $p<0.020$].

Correlations between behavioral and FFR changes

Another way to look at the effects of training on the FFR is to assess whether changes in behavioral performance correlate with changes in FFR strength. We assessed the correlation between the threshold
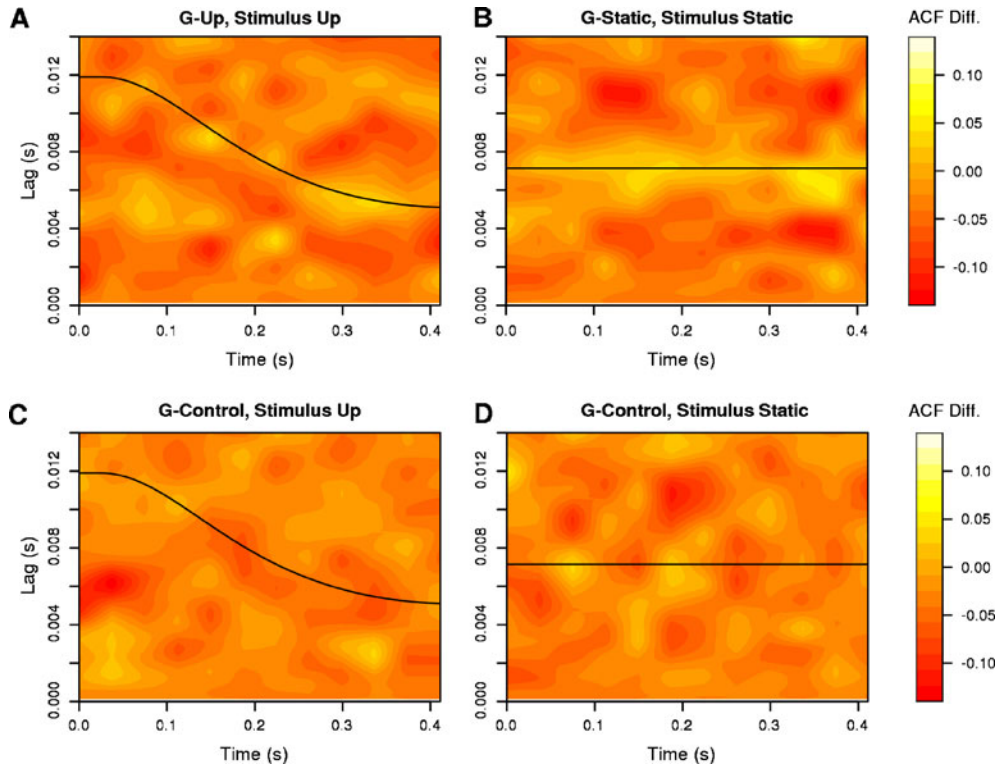
FIG. 7. Difference grand average autocorrelograms between the post and pre-FFR assessment sessions. For the group trained with the rising pitch tone (**A**), and the group trained with the static pitch tone (**B**) lighter bands, indicating greater ACF values at the second FFR session, can be seen around lags corresponding to the inverse of the F0 trajectory of the stimuli, indicated by the black line. For the control group (**C**), **D** changes of the ACF values are diffuse and do not generally correspond to the inverse of the F0 trajectory of the stimuli.
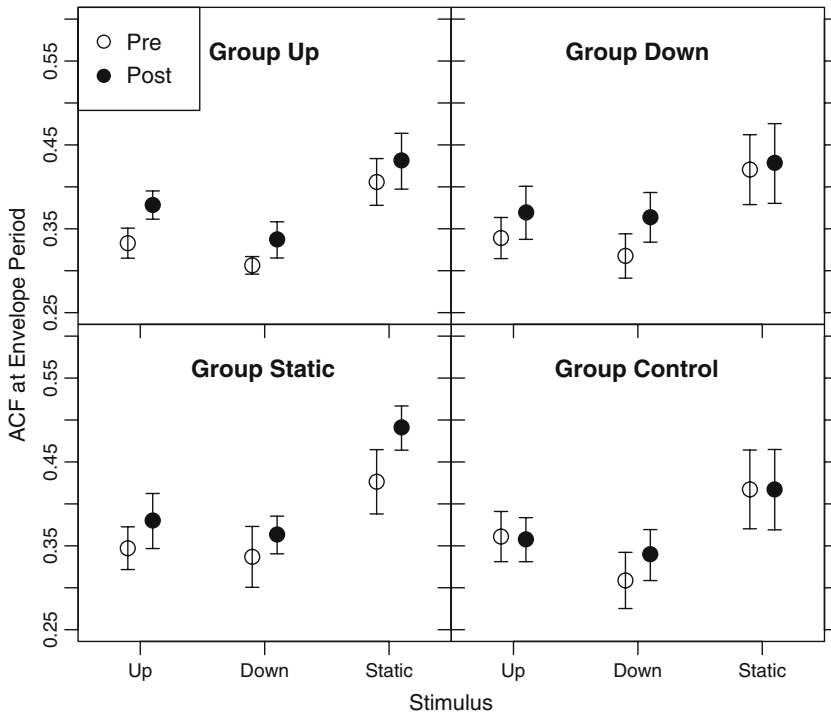

FIG. 8. ACF values at the time lags corresponding to the inverse of the F0 of the stimuli at the pre- and post-FFR assessment sessions for each stimulus and group.

changes and the $ACF_{1/F0}$ changes averaged across stimuli, including the participants from all groups in the analysis. Figure 9 displays the relationship between these two variables graphically. Their correlation was marginally non significant [Spearman's $\rho=0.275$, $p=0.055$]. When looking at each stimulus individually, changes in FFR strength significantly correlated with changes in threshold for the S-Up [$\rho=0.465$, $p<0.003$] and S-Static [$\rho=0.383$, $p<0.012$], but not for the S-Down stimulus [$\rho=0.042$, $p=0.404$].

DISCUSSION

We found changes in subcortical electophysiological responses to sounds after a multiple-hour period of pitch discrimination training in adults. These results provide direct evidence of short-term subcortical plasticity in adults. This plasticity consisted of more robust phase locking of the FFR to the static or dynamic F0 of the trained stimuli. More robust FFR phase locking to the F0 of a stimulus can reflect either a greater accuracy of phase locking of single fibers to the F0 of the stimulus, or a greater proportion of fibers phase locking to the stimulus period. The latter may be achieved either through recruitment of additional fibers phase locking to the stimulus period or through the inhibition of fibers firing at different periods. Inhibitory and excitatory circuits local to the brainstem (Yang and Pollak 1997; Burger and Pollak 1998) may mediate such changes in phase locking selectivity.

The specificity of the FFR enhancements we found with respect to pitch shape (dynamic vs. static) suggests that different mechanisms may be affected by training with static and dynamic pitch contours. FFR enhancements in Mandarin speakers are greater
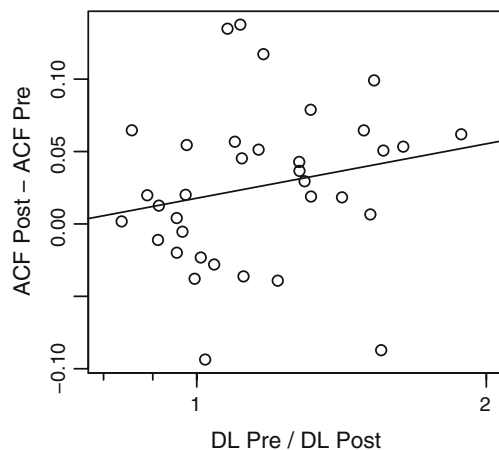

FIG. 9. Scatterplot showing the relationship between changes in behavioral thresholds and FFR strength, averaged across stimuli. The *line* represents a linear regression fit.

for tonal segments with a dynamic pitch contour (Krishnan et al. 2009b; Swaminathan et al. 2008). The identification of Mandarin tones requires the discrimination between different shapes of F0 contours rather than the discrimination of static F0 contours differing in F0 height. Our results are consistent with the idea that long-term practice with dynamic pitch stimuli in Mandarin speakers may affect FFR mechanisms specific to dynamic pitch contours (Krishnan and Gandour 2009). We did not find evidence, however, that for dynamic pitch stimuli, FFR enhancements are specific to the pitch trajectory (specificity for the rising vs. falling stimulus). FFR enhancement specificity for particular pitch trajectories has been previously found in Mandarin speakers (Krishnan et al. 2009a). Interestingly, the present behavioral results showed specificity of learning with respect to pitch trajectory. This suggests that the representation of the rising and falling pitch stimuli was differentially affected by learning at a higher processing level than the one probed by the FFR. It is possible that such specificity of learning in high-level stimulus representations guides, during long-term learning, the specificity of FFR plasticity observed in Mandarin speakers.

We also found that changes in behavioral performance in the pitch discrimination tasks correlated with changes in FFR strength for the stimuli with a rising and static pitch contour. These correlations suggest that increases in FFR strength may contribute directly to improvements in the perception of the stimuli. The proportion of variance in the changes in behavioral thresholds explained by the changes in FFR strength was relatively small. The fact that a number of participants performed close to floor level in the first behavioral threshold assessment session may have reduced the strength of the observed correlations. Nonetheless, it is likely that improvements in a perceptual discrimination task reflect improvements in stimulus encoding at several levels of sensory processing (Ahissar and Hochstein 2004). Moreover, improved perceptual discrimination may reflect improvements of other perceptual processes involved in the discrimination task, such as attentional selection of task relevant information (Goldstone 1998; Amitay 2009). The results of our study suggest that improvements in the encoding of the stimuli at the level of the brainstem make a small but significant contribution to short-term pitch discrimination learning.

Subcortical plasticity in the auditory system

The sensitivity of the FFR to multiple-hour auditory discrimination training shows that the human auditory system is susceptible to plasticity at a relatively peripheral level of sensory processing even in adult-

hood. Previous reports of plasticity in the adult human auditory system have been generally limited to cortical measures of auditory processing. There is, however, increasing evidence that short-term training (de Boer and Thornton 2008; Song et al. 2008) or short-term sensory deprivation (Munro and Blount 2009) can modify subcortical measures of auditory processing even in adulthood. The results of our study complement a growing body of research indicating that long-term experience in the discrimination and identification of pitch-evoking stimuli, obtained through the acquisition of a tone language or musical practice, modifies auditory processing at the level of the brainstem (see Tzounopoulos and Kraus 2009; Krishnan and Gandour 2009 for reviews). These studies have shown that speakers of tone languages (Krishnan et al. 2005, 2010) and musicians (Wong et al. 2007; Bidelman et al. 2009) have enhanced subcortical phase locking to the envelope of periodic sounds in comparison to English speakers without musical experience. In speakers of a tone language, these effects are present for both speech-like and non-speech stimuli, and are greater for tonal segments with highly accelerated F0 contours (Xu et al. 2006; Swaminathan et al. 2008; Krishnan et al. 2009b), which are characteristic of tone languages (Eady 1982). These effects have been measured in native speakers of a tone language, and musicians who started practicing during childhood. It is known that certain forms of neural plasticity are possible only during limited critical developmental periods at a young age (Hensch 2004). These include the postnatal reorganization of tonotopic maps in the auditory cortex of rats (Chang and Merzenich 2003; de Villers-Sidani et al. 2007), the alignment of auditory space maps in the barn owl midbrain to altered visual space maps (Knudsen et al. 2000), and the proficient acquisition of language after cochlear implantation in humans (Manrique et al. 1999; Harrison et al. 2005). The results of our study show that FFR plasticity is not limited to a critical period during childhood, although it is still possible that FFR plasticity is greater during early development. The larger size of the FFR enhancement effects observed in speakers of a tone language (e.g., Krishnan et al. 2005) compared to the present study is compatible with this hypothesis. However these differences may also be due to differences in the amount of "training", which can be measured in terms of years for native tone language speakers, and only hours for the participants of our study. The correlations found by Wong et al. (2007) between FFR measures of pitch processing for some tones, and years of musical training in musicians, suggest that although experience-dependent FFR enhancements effects can be measured only after a few hours

of training, they build-up progressively with further training.

In principle, neurophysiological differences between different categories of listeners may also be explained by factors other than neural plasticity (Monaghan et al. 1998). The results of our study suggest that neural plasticity is likely to be the cause of the FFR enhancements in "pitch experts" (musicians and speakers of a tone language) that have been reported by previous studies. A comprehensive understanding of interindividual differences in pitch-related abilities and their neurophysiological correlates, however, requires taking into account the possible contribution of other factors to these differences, such as genetic predispositions. It has been shown, for example, that the ability to identify incorrect tones in familiar melodies is a highly heritable trait, with genetic factors explaining up to 80% of variability in this ability (Drayna et al. 2001). Moreover, there is evidence that 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). Studies comparing pitch-processing measures between different populations of listeners cannot disentangle the contribution of neural plasticity from the contribution of genetic factors to the differences measured. Studies comparing pitch-processing measures before and after a period of auditory training, as well as studies comparing pitch-processing measures between populations which are unlikely to have systematic genetic differences, but differ in their experience with pitch discrimination (e.g., Chinese Mandarin speakers vs. Chinese English speakers) can best identify the contribution of neural plasticity to pitch processing abilities.

The exact mechanisms underlying the plasticity of the FFR response remain currently unknown. In a recent study, de Boer and Thornton (2008) found that activity of the medial olivocochlear bundle, which sends efferent signals from the brainstem to the cochlea, and is part of the corticofugal efferent system, reflected improvements on a speech in noise discrimination task after a period of auditory discrimination training. The corticofugal system, that projects from the auditory cortex to all major brainstem nuclei (Winer 2006), is likely to play a crucial role in subcortical plasticity. Bajo et al. (2010) have shown that ferrets subject to monaural sensory deprivation cannot relearn to localize sound accurately after damage to the corticocollicular pathway, which is part of the corticofugal system. Studies in other non-human species have demonstrated short-term changes in frequency selectivity of neurons in the inferior colliculus and cochlear nucleus after a period of auditory fear conditioning or focal electrical stimula-

tion of the auditory cortex (Suga and Ma 2003; Suga 2008; Luo et al. 2008). Subcortical plasticty elicited by electrical stimulation of the auditory cortex can be explained only by the activation of the corticofugal system. It is possible that the subcortical plasticity observed in the present experiment depends on similar tuning of F0 specificity in neurons under the influence of descending projections.

ACKNOWLEDGMENTS

REFERENCES

AHISSAR M, HOCHSTEIN S (2004) The reverse hierarchy theory of visual perceptual learning. Trends Cogn Sci 8:457–464

AMITAY S (2009) Forward and reverse hierarchies in auditory perceptual learning. Learn Percept 1:59–68

ANANTHANARAYAN AK, DURRANT JD (1992) The frequency-following response and the onset response: evaluation of frequency specificity using a forward-masking paradigm. Ear Hear 13:228–232

BAJO VM, NODAL FR, MOORE DR, KING AJ (2010) The descending corticocollicular pathway mediates learning-induced auditory plasticity. Nat Neurosci 13:253–260

BIDELMAN GM, GANDOUR JT, KRISHNAN A. (2009) Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J Cogn Neurosci (in press)

DE BOER J, THORNTON ARD (2008) Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. J Neurosci 28:4929–4937

BURGER RM, POLLAK GD (1998) Analysis of the role of inhibition in shaping responses to sinusoidally amplitude-modulated signals in the inferior colliculus. J Neurophysiol 80:1686–1701

CHANG EF, MERZENICH MM (2003) Environmental noise retards auditory cortical development. Science 300:498–502

CLARK JL, MOUSHEGIAN G, RUPERT AL (1997) Interaural time effects on the frequency-following response. J Am Acad Audiol 8:308–313

CONOVER W, JOHNSON M, JOHNSON MA (1981) Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data. Technometrics 23:351–361

DAU T (2003) The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. J Acoust Soc Am 113:936–950

DEDIU D, LADD DR (2007) Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. Proc Natl Acad Sci USA 104:10944–10949

DRAYNA D, MANICHAIKUL A, DE LANGE M, SNIEDER H, SPECTOR T (2001) Genetic correlates of musical pitch recognition in humans. Science 291:1969–1972

EADY SJ (1982) Differences in the F0 patterns of speech: tone language versus stress language. Lang Speech 25:29–42

GARDI J, MERZENICH M, MCKEAN C (1979) Origins of the scalp recorded frequency-following response in the cat. Audiology 18:358–381

GOLDSTONE RL (1998) Perceptual learning. Annu Rev Psychol 49:585–612

GRIMAULT N, MICHEYL C, CARLYON RP, COLLET L (2002) Evidence for two pitch encoding mechanisms using a selective auditory training paradigm. Percept Psychophys 64:189–197

HARRISON RV, GORDON KA, MOUNT RJ (2005) Is there a critical period for cochlear implantation in congenitally deaf children? Analyses of hearing and speech perception performance after implantation. Dev Psychobiol 46:252–261

HENSCH TK (2004) Critical period regulation. Annu Rev Neurosci 27:549–579

KNUDSEN EI, ZHENG W, DEBELLO WM (2000) Traces of learning in the auditory localization pathway. Proc Natl Acad Sci USA 97:11815–11820

KRISHNAN A, GANDOUR JT (2009) The role of the auditory brainstem in processing linguistically-relevant pitch patterns. Brain Lang 110:135–148

KRISHNAN A, XU Y, GANDOUR J, CARIANI P (2005) Encoding of pitch in the human brainstem is sensitive to language experience. Brain Res Cogn Brain Res 25:161–168

KRISHNAN A, GANDOUR JT, BIDELMAN GM, SWAMINATHAN J (2009A) Experience-dependent neural representation of dynamic pitch in the brainstem. NeuroReport 20:408–413

KRISHNAN A, SWAMINATHAN J, GANDOUR JT (2009B) Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. J Cogn Neurosci 21:1092–1105

KRISHNAN A, GANDOUR JT, BIDELMAN GM (2010) The effects of tone language experience on pitch processing in the brainstem. J Neurolinguist 23:81–95

LEVITT H (1971) Transformed up-down methods in psychoacoustics. J Acoust Soc Am 49(2):467+

LIU L, PALMER AR, WALLACE MN (2006) Phase-locked responses to pure tones in the inferior colliculus. J Neurophysiol 95:1926–1935

LUO F, WANG Q, KASHANI A, YAN J (2008) Corticofugal modulation of initial sound processing in the brain. J Neurosci 28:11615–11621

MANRIQUE M, CERVERA-PAZ FJ, HUARTE A, PEREZ N, MOLINA M, GARCÍA-TAPIA R (1999) Cerebral auditory plasticity and cochlear implants. Int J Pediatr Otorhinolaryngol 49(1):S193–197

MCLACHLAN N (2009) A computational model of human pitch strength and height judgments. Hear Res 249:23–35

MEDDIS R, O'MARD LP (2006) Virtual pitch in a computational physiological model. J Acoust Soc Am 120:3861–3869

MONAGHAN P, METCALFE NB, RUXTON GD (1998) Does practice shape the brain? Nature 394:434

MUNRO KJ, BLOUNT J (2009) Adaptive plasticity in brainstem of adult listeners following earplug-induced deprivation. J Acoust Soc Am 126:568–571

PLACK CJ, OXENHAM AJ (2005) The present and future of pitch. In: Plack CJ, Oxenham AJ (eds) Pitch: neural coding and perception. Springer, New York

PLACK CJ, OXENHAM AJ (2005b) The Psychophysics of Pitch. In Plack CJ, Oxenham AJ (eds) Pitch: neural coding and perception Springer, New York

SATTERTHWAITE FE (1946) An Approximate Distribution of Estimates of Variance Components. Biometrics Bulletin 2:110–114

SMITH JC, MARSH JT, BROWN WS (1975) Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. Electroencephalogr Clin Neurophysiol 39:465–472

SONG JH, SKOE E, WONG PCM, KRAUS N (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. J Cogn Neurosci 20:1892–1902

SUGA N (2008) Role of corticofugal feedback in hearing. J Comp Physiol A Sens Neural behav physiol 194:169–183

SUGA N, MA X (2003) Multiparametric corticofugal modulation and plasticity in the auditory system. Nat Rev Neurosci 4:783–794

Swaminathan J, Krishnan A, Gandour JT (2008) Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. NeuroReport 19:1163–1167

Tzounopoulos T, Kraus N (2009) Learning to encode timing: mechanisms of plasticity in the auditory brainstem. Neuron 62:463–469

de Villers-Sidani E, Chang EF, Bao S, Merzenich MM (2007) Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. J Neurosci 27:180–189

Winer JA (2006) Decoding the auditory corticofugal systems. Hear Res 212:1–8

Wong PCM, Skoe E, Russo NM, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci 10:420–422

Xu Y, Krishnan A, Gandour JT (2006) Specificity of experience-dependent pitch representation in the brainstem. NeuroReport 17:1601–1605

Yang L, Pollak GD (1997) Differential response properties to amplitude-modulated signals in the dorsal nucleus of the lateral lemniscus of the mustache bat and the roles of GABAergic inhibition. J Neurophysiol 77:324–340