Title: Getting back on the beat: links between auditory-motor integration and precise auditory processing at fast time scales

Running title: Neural foundations of auditory-motor adaptation

Keywords: auditory, motor, adaptation, synchronization

Journal section: Cognitive neuroscience

Authors: Adam Tierney[1,2,6,7], Nina Kraus[1-6†]

Affiliations

[1]Auditory Neuroscience Laboratory, [2]Department of Communication Sciences, [3]Institute for Neuroscience, [4]Department of Neurobiology and Physiology, [5]Department of Otolaryngology, [6]Northwestern University, Evanston, Illinois USA, [7]Now at: Department of Psychological Sciences, Birkbeck, University of London, London, UK

†Corresponding author
Nina Kraus, Ph.D.
Northwestern University
2240 Campus Drive
Evanston IL, 60208
USA
Email: nkraus@northwestern.edu
Phone: 847-491-3181
Fax: 847-491-2523

Number of pages: 17

Number of figures: 4

Words in abstract: 250

Words in the introduction: 559

Words in the manuscript: 5577

Abstract

The auditory system is unique in its ability to precisely detect the timing of perceptual events and use this information to update motor plans, a skill crucial for language. The characteristics of the auditory system which enable this temporal precision, however, are only beginning to be understood. Previous work has shown that participants who can tap consistently to a metronome have neural responses to sound with greater phase coherence from trial to trial. We hypothesized that this relationship is driven by a link between the updating of motor output by auditory feedback and neural precision. Moreover, we hypothesized that neural phase coherence at both fast time scales (reflecting subcortical processing) and slow time scales (reflecting cortical processing) would be linked to auditory-motor timing integration. To test these hypotheses we asked participants to synchronize to a pacing stimulus and then changed either the tempo or the timing of the stimulus to assess whether they could rapidly adapt. Participants who could rapidly and accurately resume synchronization had neural responses to sound with greater phase coherence. However, this precise timing was limited to the time scale of 10 ms (100 Hz) or faster; neural phase coherence at slower time scales was unrelated to performance on this task. Auditory-motor adaptation, therefore, specifically depends upon consistent auditory processing at fast, but not slow, time scales.

Introduction

The auditory system precisely tracks the timing of perceptual events and coordinates this information with the motor system. Studies of synchronized tapping, for example, have shown that participants are less variable when tapping to an auditory beat than a visual beat (Kolers and Brewster 1985, Chen et al. 2002, Patel et al. 2005, Hove et al. 2013). This precise auditory-motor integration may also be important for auditory perception, as listeners may categorize speech sounds and track temporal patterns by mapping acoustic information onto articulatory gestures (Kotz and Schwartze 2010, Schulze et al. 2012, Peelle and Davis 2012). However, little is known about the mechanisms that make this auditory-motor precision possible.

One way to investigate the neural foundations of precise auditory-motor timing integration is to examine individual differences among participants. Participants whose frequency-following neural response to sound displayed greater phase coherence from trial to trial tap more consistently to a beat (Tierney and Kraus 2013) and are better able to read (Hornickel and Kraus 2013). This suggests that to perform auditory-motor temporal integration, whether when acquiring language skills or performing a non-linguistic rhythm task, the auditory system must consistently represent the timing of sounds on a millisecond-level timescale. However, synchronizing to a metronome is a complex process consisting of several components, including the production of steady movement, reliable internal time-keeping, and adaptation to changes in the timing relationship between sound and movement (Semjen et al. 1998, Thaut et al. 1998). It remains unknown, therefore, which of these components relies upon consistent auditory processing. In particular, it remains to be demonstrated that auditory-motor integration is specifically linked to auditory neural precision. *We hypothesized that auditory neural temporal precision is vital for auditory-motor temporal integration. To test this hypothesis we predicted that millisecond-level auditory neural phase coherence would be linked to the ability to adapt to small shifts in the timing of sound events during synchronization.*

Another open question is how auditory-motor temporal integration depends upon neural precision at different time scales. Previous work reported a relationship between the precision of high-

frequency auditory responses (10 ms and faster) and synchronization performance (Tierney and Kraus 2013). Neural precision at time scales of 10 ms and faster (100 Hz and above) may be particularly vital because the sound-to-movement alignment errors that arise during synchronization occur at the level of a few milliseconds. Precision at slower time scales, however, may also be important, given that low-frequency neural entrainment tracks the rhythmic structure of sound sequences (Large 2008, Nozaradan et al. 2011, 2012, Tierney et al. 2014). *Given an emerging view of the auditory system as a distributed, but integrated, neural circuit (Kraus & White-Schwoch in press) we hypothesized that neural precision across both slow and fast time scales is important for accurate auditory time perception. To test this hypothesis, we predicted that phase coherence across multiple time scales would be linked to auditory-motor temporal integration ability.* Here we investigated the neural correlates of auditory-motor timing integration skill by asking participants to synchronize to a metronome that changed either its tempo or its timing. Neural responses to speech were collected to determine how temporal adaptation ability related to neural phase coherence at the faster time scale of the frequency-following response (10 ms or less, corresponding to 100 Hz and above) and the slower time scale of the evoked cortical response (100-200 ms, corresponding to 5-10 Hz).

Methods

Subjects

65 participants (31 female) were recruited from local high schools and universities. Mean participant age was 18.10 years (standard deviation 0.98). Participants had no diagnosis of a learning or neurological disorder and had normal hearing thresholds less than 20 dB (re: normal hearing level) for octaves between 250 and 8000 Hz. Click-evoked auditory brainstem response latencies were within normal limits, as established by our laboratory (5.41-5.97 ms; 100-µs click stimulus presented at 80 dB sound pressure level (SPL) at a rate of 31/s). For participants younger than age 18, informed assent was obtained from the participant and consent from the participant's legal guardian. Informed consent was obtained for all participants older than age 18. All procedures were approved by the Institutional Review Board at Northwestern University.

Drumming

Apparatus

Stimuli for the synchronization tests were created in Matlab (The Mathworks), stored on an iPod Nano (Apple), and presented over headphones. The drum sound used in both tests was a 150-ms recording of a struck conga drum acquired at freesound.org. Participants listened to the stimulus track while drumming on a conga drum with a drum trigger (Pulse Percussion) placed against the underside of the drum head. The output of the drum trigger was combined with a copy of the auditory signal being presented to the participant as two channels of a stereo input to a computer running the audio recording program Audacity. Thus, the timing relationship between the stimulus input and the participant's drumming was captured in real time.

Processing

Drumming data were processed by a custom-made program which took as input the audio file created in Audacity and returned a list of stimulus presentation and drum onset times. Continuous variations in amplitude were transformed into discrete onsets by setting an amplitude threshold and a

relaxation time; time points were marked as onsets if 1) the amplitude of the point exceeded the amplitude threshold, and 2) a length of time equal to or greater than the relaxation time had passed since the last time point when the threshold was exceeded. Amplitude thresholds and relaxation times were set manually on a subject-by-subject basis to account for inter-individual variation in the exact manner in which the drum was struck. Drum and stimulus onsets were checked by visually comparing a vector of onset times to the raw amplitude contour, to ensure that all onsets were correctly marked and that all marked onsets corresponded to a drum or stimulus onset. These stimulus and drum onsets were then subjected to further processing to produce the summary scores for each test, as described below.

Tempo adaptation test

This test measured participants' ability to rapidly adapt to a shift in a metronome's tempo. Participants were presented with fifty-five trials. Each trial consisted of between 11 and 15 presentations of a conga drum sound. In each trial the first 6 to 10 sounds were separated by 500-ms intervals. During five of the fifty-five trials the last five sound presentations continued at the same rate. The other fifty trials included a tempo shift: in twenty-five of the trials the stimulus rate switched to a slower tempo for the last five sound presentations, with five trials each using tempos of 550, 540, 530, 520, and 510 ms. In the remaining twenty-five trials the stimulus rate switched to a faster tempo for the last five presentations, with five trials each using tempos of 450, 460, 470, 480, and 490 ms. Condition order was pseudorandomized so that both the degree and direction of the tempo shift (i.e. faster or slower) was unpredictable. This test lasted approximately ten minutes.

Participants were asked to drum along to the beat. They were told that at some point the drum beat might change to a new tempo, and that if they heard a tempo shift they should switch the tempo of their drumming to match the new drum tempo. Performance was assessed for each trial by measuring the difference between the target tempo and each of the last two intervals produced. For example, if a given trial's target tempo was 460 ms and the last two intervals the participant produced were 450 and 480 ms the participant's score for that trial would be (10 + 20)/2 or 15. Lower scores indicated better performance.

Timing shift adaptation test

This test measured participants' ability to resume synchronization with a metronome after a single metronome interval was lengthened or shortened. A total of four trials were presented to the participants. During each trial participants heard a conga sound presented with a constant inter-stimulus interval of 500 ms and they were asked to synchronize to the beat. Participants were told that one of the sounds might occasionally sound as if it were off of the beat, but that they should do their best to keep their movements in time with the sounds they were hearing. Occasionally a single interval was either lengthened or shortened. Each trial contained a total of 16 shifted beats that were separated by eight isochronous beats to give the participant time to resume synchronization. In two of the four trials, the interval was lengthened or shortened by 10 ms, while in the remaining two trials, the interval was lengthened or shortened by 50 ms. Order of presentation of lengthened versus shortened intervals was pseudorandomized so that the direction of the shift was unpredictable.

This single altered time interval increased the offset between the participants' drumming and the stimulus. To minimize this offset, participants needed to resynchronize to the beat. This test, therefore, provides an efficient way to assess participants' ability to rapidly synchronize. The degree of

synchronization following an altered metronome interval was quantified as the standard deviation of the offset between the participant's movements and the drum sound onset for the six beats following the shifted interval. Performance was averaged across all four trials. Lower scores indicated better performance.

Electrophysiology

Stimulus and recording

The stimulus for the neurophysiological testing was a 170-ms 'da' synthesized using a Klatt synthesizer. The stimulus consisted of a 5-ms onset burst followed by a 45 ms consonant-vowel transition period followed by a 120 ms steady-state vowel period. From 5 to 50 ms the fundamental frequency stayed constant at 100 Hz, the first formant increased from 400 to 720 Hz, the second formant decreased from 1700 to 1240 Hz, and the third formant decreased from 2580 to 2500 Hz. From 50 to 170 ms the fundamental frequency remained constant at 100 Hz while the first, second, and third formants stayed constant at 720, 1240, and 2500 Hz, respectively. The fourth, fifth, and sixth formants stayed constant from 5 to 170 ms at 3300, 3750, and 4900 Hz, respectively. Stimuli were presented using Neuroscan Stim[2] (Compumedics) at 80 dB SPL at a rate of 3.95 Hz through insert earphones (ER-3; Etymotic Research). Stimuli were presented in alternating polarities; adding neural responses to alternating polarities ensured that the neural measures described below were not contaminated by stimulus artifact. Presenting stimuli in alternating polarities is a standard way of accentuating both the phase coherence and amplitude of the F0 response, as adding the alternate polarities accentuates the lower-frequency response envelope, while subtracting the polarities accentuates the higher-frequency fine structure (Aiken and Picton 2008). Electrophysiological data were collected using Neuroscan Acquire 4.3. A montage of three Ag-AgCl electrodes was used with the active electrode at Cz, ground at the forehead, and reference at the right earlobe. All impedances were kept below 5 kΩ, with inter-electrode differences kept below 3 kΩ. Data were digitally sampled at 20000 Hz. During the recording session participants sat quietly while watching a movie to maintain a state of calm alertness. 6000 artifact-free trials were collected (3000 of each polarity).

Stimuli were presented in two different conditions. In the Quiet condition, stimuli were presented in the absence of background noise. A Noise condition was also presented to determine whether stressing the auditory system via background noise would elicit a wider range of inter-subject phase-locking, potentially strengthening links between neural function and auditory-motor precision. In the Noise condition, therefore, stimuli were presented in the context of multi-talker background babble at a signal-to-noise ratio of +10 dB relative to the root-mean-square amplitude of the noise. Both stimulus and noise were presented monaurally to the right ear.

EEG data processing

The phase coherence of the neural response to sound was examined by calculating inter-trial phase-locking. Raw neural data were epoched from -40 to 250 ms around stimulus onset. Trials with amplitude exceeding 35 microvolts were rejected as artifact. Epochs were baseline corrected to the mean amplitude of the epoch and detrended, and the first and final 15 ms of each epoch were ramped with a Hann window. A set of complex Morlet wavelets was then convoluted with each response trial at each time point, providing a measure of the amplitude and phase for each time x frequency bin. To analyze the inter-trial phase-locking of the frequency-following response, complex Morlet wavelets with

center frequencies of 70 to 410 Hz in 1 Hz steps were used, with the bandwidth parameter set as 6.5 Hz. (Setting the bandwidth parameter to a constant value results in a constant window duration being applied across frequencies.) To analyze the inter-trial phase-locking of the low-frequency evoked response, wavelets with center frequencies of 1 to 50 Hz in 1 Hz steps were used, with the bandwidth parameter set as 0.5 Hz. Amplitude provides a measure of the amount of energy present in the signal at that frequency, while phase provides an estimate of the timing of the response at that frequency. Averaging normalized phase vectors across trials and then calculating the length of the resulting vector, therefore, gives a measure of the extent to which the response varies in phase at that time x frequency point. A phase-locking value of 1 would indicate perfect inter-trial synchrony, while a small phase-locking value would indicate no inter-trial neural synchrony. Phase-locking data were displayed in 3D color plots (see Figure 2). For statistical analysis, frequency-following response phase-locking was quantified as average phase-locking in 20 Hz bins centered around the fundamental frequency and the first through third harmonics (100, 200, 300, and 400 Hz) from 5 to 170 ms. Visual analysis revealed phase-locking below 50 Hz was primarily limited to the high theta range (5-10 Hz). Low-frequency evoked response phase-locking was, therefore, calculated as average phase-locking between 5 and 10 Hz from 5 to 170 ms.

As a confirmatory measure the consistency of the neural response was also calculated (Hornickel et al. 2012). First, raw continuous data were band-passed using one of two pass bands: low-frequency (1-50 Hz) or high-frequency (70-2000 Hz). Data were then epoched from -40 to 210 ms and trials with amplitude exceeding +/- 35 microvolts were rejected as artifact. Epochs were then baselined to the amplitude of the prestimulus period. A total of 6000 artifact-free trials were analyzed. A random sampling process was repeated 300 times to produce an estimate of the average consistency between responses to the stimulus. For each sample, 3000 randomly-selected trials were averaged, and then the remaining 3000 trials were combined to form a second average. The portions of these two waveforms between 5 and 170 ms after sound onset were then correlated. A response that is more consistent from trial to trial will lead to highly similar subaverages and a more robust correlation, while more variability from trial to trial will lead to dissimilar subaverages and a weaker correlation. Finally, the 300 resulting r-values produced by this process were averaged and converted to a z-score via a Fisher transform to create a global response consistency score.

Statistical analysis

To confirm that the relationship between phase-locking and auditory-motor adaptation held across all tempo shift conditions a repeated measures ANOVA was run with phase-locking group (top versus bottom third) as the between-subjects factor, tempo shift condition as the within subjects factor, and adaptation performance as the dependent variable. To determine whether the addition of noise impacted the relationship between auditory-motor adaptation ability and inter-trial phase coherence, a series of repeated measures ANOVAs were run with drumming performance group as the between subjects variable (top versus bottom third performers) and stimulus presentation condition (quiet vs noise) as the within subjects variable. Separate ANOVAs were run for the timing shift versus tempo adaptation tests and for high-frequency versus low-frequency phase-locking, for a total of four ANOVAs.

The following data were normally distributed, and were therefore not transformed: composite data for the period correction and timing shift correction tests, high-frequency and low-frequency phase-locking in quiet and noise. Response consistency data, however, were not normally distributed,

and therefore outliers of greater than two standard deviations away from the mean were removed to achieve normality. The critical phase locking value at p = 0.05 was calculated as 0.02234 using the estimation of the circular mean described in Fisher (1996). Six subjects showed frequency-following response phase-locking values for F0 through H4 (100, 200, 300, and 400 Hz) that were not significant according to this analysis, and were excluded from our main analysis.

To determine the relationship between auditory-motor integration ability and neural precision, four stepwise linear regressions were run. In one, low-frequency evoked and frequency following response phase coherence were included as possible predictors of tempo adaptation. In another, low-versus high-frequency phase coherence predicted phase adaptation. In the remaining two, low-frequency and high-frequency response consistency predicted tempo and phase adaptation. Follow-up Spearman correlational analyses were run comparing performance on each synchronization test to composite phase-locking at high and low frequencies averaged across quiet and noise conditions, as well as to composite response consistency of responses band-passed with low and high pass bands. Spearman correlational analyses were used to further investigate relationships between inter-trial phase-locking, response consistency, and adaptation ability. Correlation confidence intervals were calculated using the method of Bonett and Wright (2000).

Results

There was no interaction between phase-locking group and tempo shift (F(1,42) = 1.61, p = 0.149), confirming that the degree of shift did not affect the relationship between auditory motor adaptation. Similarly, for timing shifts, there was no interaction between phase-locking group and degree of shift (F(1,42) = 2.01, p = 0.129). Therefore we collapsed across all tempo and timing shift conditions prior to our main analysis.

Participants showed greater phase-locking to sounds presented in quiet, compared to in noise (main effect of recording condition (quiet vs. noise; F(1,42) = 42.85, p < 0.001). However, the extent to which phase-locking was enhanced in the participants who were better able to perform tempo adaptation did not differ between the quiet and noise conditions (no interaction between group and recording condition, F(1,42) = 0.260, p = 0.613). Similarly, for the timing shift adaptation test, the extent to which phase-locking was enhanced in the good adaptors compared to the poor adaptors did not differ between quiet and noise conditions (no interaction between group and recording condition, F(1,42) = 1.272, p = 0.266). For the analyses in the remainder of the paper, therefore, we produced composite measures by averaging phase coherence and response consistency across quiet and noise conditions.

Auditory-motor adaptation and phase coherence

See Figure 1 for a histogram of phase coherence across F0 through H4 (not including the six subjects who were excluded for having phase coherence values falling below the critical value). Across responses to speech presented in both quiet and noise, inter-trial phase-locking at frequencies corresponding to the fundamental frequency and first three harmonics of the stimulus (90-110, 190-210, 290-310, and 390-410 Hz) was higher in participants who were better at adapting to perturbation while synchronizing (Fig. 2 left; Figs. 3-4). However, across responses to speech presented in both quiet and noise, inter-trial phase-locking at the frequencies that characterize the cortical auditory evoked response (5-10 Hz) was not linked to the ability to adapt to perturbation while synchronizing (Fig. 2,

right; Fig. 3). Stepwise linear regression revealed that high-frequency phase coherence predicted tempo adaptation performance (B = -187.457 (confidence interval -297.329 to -77.584), R^2 = 0.170, p = 0.001), but low-frequency phase coherence did not significantly improve the model (p = 0.694). Similarly, high-frequency phase coherence predicted timing shift adaptation performance (B = -151.229 (confidence interval -254.223 to -48.234), R^2 = 0.132, p = 0.005), but low-frequency phase coherence did not significantly improve the model (p = 0.783).

Auditory-motor adaptation and response consistency

Across responses to speech presented in both quiet and noise, the consistency of the response filtered at frequencies corresponding to the frequency-following response (70-2000 Hz) was higher in participants who were better at adapting to perturbation while synchronizing (Fig. 5). However, across responses to speech presented in both quiet and noise, the consistency of the response filtered at frequencies corresponding to the evoked response (1-50 Hz) was not linked to the ability to adapt to perturbation while synchronizing. Stepwise linear regression revealed that frequency-following response consistency predicted tempo adaptation performance (B = -10.157 (confidence interval -15.612 + -4.703), R^2 = 0.219, p < 0.001), but low-frequency evoked response consistency did not significantly improve the model (p = 0.716). Similarly, frequency-following response consistency tended to predict timing shift adaptation performance (B = -4.774 (confidence interval -10.392 to 0.845), R^2 = 0.055, p = 0.094), but low-frequency evoked response consistency did not significantly improve the model (p = 0.817).

Correlational analyses

Correlational analyses were used to further investigate relationships between inter-trial phase-locking, response consistency, and adaptation ability (Figs. 5-7). Given the lack of an interaction between group and recording condition in the analyses reported above, a composite phase-locking score was generated by averaging values for the quiet and noise responses. Subjects who performed better on the tempo adaptation task had stronger phase-locking (rho = -0.416, p = 0.001, confidence interval -0.614 to -0.168) and response consistency (rho = -0.447, p < 0.001, confidence interval -0.643 to -0.196) at the frequencies characterizing the frequency-following response but did not display an advantage for phase-locking (rho = 0.068, p = 0.608, confidence interval -0.192 to 0.319) and response consistency (rho = -0.021, p = 0.881, confidence interval -0.248 to 0.287) at the lower frequencies that make up the evoked response. Similarly, phase adaptation tracked with phase-locking (rho = -0.343, p = 0.008, confidence interval -0.556 to -0.088) and response consistency (rho = -0.270, p = 0.044, confidence interval-0.500 to -0.002) at the frequencies characterizing the frequency-following response but did not track with phase-locking (rho = 0.059, p = 0.658, confidence interval -0.200 to 0.311) or response consistency (rho = 0.045, p = 0.745, confidence interval -0.226 to 0.309) at the lower frequencies which make up the evoked response. Tempo and timing shift adaptation were correlated (Fig. 6; rho = 0.477, p < 0.001, confidence interval 0.238 to 0.662). Finally, inter-trial phase-locking of the frequency following response was not correlated with phase-locking of the low-frequency evoked response (rho = 0.067, p = 0.614, confidence interval -0.193 to 0.318).

Discussion

As predicted, participants who were better able to rapidly adapt to a changing metronome tempo or more quickly resynchronize to a beat after a timing shift had neural responses with greater

phase coherence from trial to trial. However, contrary to our predictions, this greater neural precision was limited to the higher frequencies that make up the subcortical frequency-following response (90 Hz to 410 Hz, corresponding to time scales of 2.5 to 11 ms). Good and poor adapters did not differ in neural precision at the lower frequencies that characterize the cortical evoked response (5-10 Hz, or 100 to 200 ms).

Humans can accurately synchronize to sound by rapidly and precisely adapting to errors that emerge from transient fluctuations in stimulus input and motor output (Repp 2000, Madison and Merker 2004). However, not everyone is equally skilled at this auditory-motor integration, and we predicted that phase coherence of neural sound processing across slow and fast time scales would help determine error correction skill. As predicted, we found that participants whose frequency-following responses to sound displayed greater phase coherence from trial to trial were better able to adapt to timing shifts while synchronizing. However, contrary to our predictions we found that the phase coherence of the lower-frequency cortical response was unrelated to auditory-motor adaptation skills. These results suggest that auditory neural precision on the time scale of 100 Hz and above (periods of 10 ms or less) that characterizes the frequency following response is crucial for accurate auditory-motor temporal integration. Auditory neural precision on the slower scale of 5-10 Hz (100-200 ms) that characterizes the cortical response, however, does not seem to underlie accurate auditory-motor adaptation. It may instead be more important for tasks that require the integration of auditory information over time, such as extraction of the beat from a complex stimulus or encoding a sequence of durations into memory, a prediction that could be investigated by future work.

We previously reported (Tierney and Kraus. 2013) that variability during beat synchronization is lower in subjects with more precise frequency-following responses. However, beat synchronization is a complex task, and as such synchronization variability can reflect different sources of variance, including motor variability, timekeeper variability, and the use of auditory feedback to correct transient increases in the asynchrony between movement and sound. Here we present evidence that temporal precision of the frequency-following response is vital for auditory-motor timing integration. The primary neural generator of the frequency-following response (Chandrasekaran and Kraus 2010) is the inferior colliculus (IC), a region that may play a crucial role in the auditory-motor error correction processes that make synchronization possible. The IC is a neural hub, receiving afferent input from peripheral auditory structures (Kudo and Niimi 1980, Coleman and Clerici 1987) and efferent input from cortical regions (Bajo et al. 2010). The inferior colliculus is also capable of precisely phase-locking to high-frequency auditory input (Liu et al. 2006), a characteristic that causes it to be the primary generator of the frequency-following response (Warrier et al. 2011). Moreover, the IC connects directly to the cerebellum, bypassing auditory cortex (Mower et al. 1979, Hashikawa 1983, Saint Marie, 1996). The cerebellum updates motor behavior by comparing expected with actual feedback (Wolpert et al. 1998) and is involved in both auditory-motor synchronization (Molinari et al. 2007, Bijsterbosch et al. 2011) and perceptual timing (Lee et al. 2007). The cerebellum, in turn, connects to the basal ganglia, the proposed location of a beat-based timing system (Grahn and Brett 2007, Teki et al. 2011, Bartolo et al. 2014) which has been linked to a deficit in tempo adaptation (Schwartze et al. 2011). The IC's temporal precision, integrative role, and direct connection to subcortical motor structures, therefore, make it ideal for communicating auditory timing information to motor regions (Molinari et al. 2005, Warren et al. 2005, Malcolm et al. 2008), a process crucial for auditory-motor integration. Our results suggest that fine temporal precision in the IC is important for accurate auditory-motor integration. Thus, temporal

precision in the auditory midbrain may be a crucial gatekeeper helping to determine individual differences in auditory-motor timing ability.

On the other hand, the lack of relationships between phase-locking at slower time scales (100-200 ms) and adaptation to timing perturbation during synchronization suggests that low-frequency cortical temporal precision is not a crucial factor driving precise auditory-motor integration. This is surprising, given that cells in premotor cortex are tuned to duration and serial order (Merchant et al. 2013, Crowe et al. 2014), that auditory and motor/premotor cortex are functionally connected during synchronization tasks (Pollok et al. 2003, Chen et al. 2006, 2008, Krause et al. 2010b), and that phase-locking of slow cortical oscillations plays a role in tracking of rhythmic information (Nozaradan et al. 2011, 2012, Tierney et al. 2014). It is possible that auditory cortex is responsible for integrating timing information over a longer time scale, on the order of hundreds of milliseconds or more, which would be more important for tasks such as remembering rhythmic sequences or predicting the timing of future rhythmic events (Merchant et al. 2015) than for synchronization. Thus, while auditory and premotor cortex do interact during synchronization, slow auditory phase-locking may not be a bottleneck for synchronization skill because beat synchronization does not rely upon integration of information across time periods longer than tens of milliseconds. Alternately, it could be that other cortical functions besides low-frequency phase-locking are important for error correction in synchronization. Perhaps, for example, subcortical high-frequency phase coherence can affect the temporal consistency of the middle latency response, which then influences synchronization via connections to pre-motor cortex.

Another possible explanation for the lack of a connection between low-frequency neural phase coherence and synchronization adaptation is that the sounds to which participants were synchronizing were characterized by very rapid amplitude onsets. For sounds with gradually ramping onsets, amplitude rise time helps to determine perceived onset time (Caclin et al. 2005). As tracking of amplitude envelope is tied to phase-locking of slow cortical oscillations (Luo and Poeppel 2007, Abrams et al. 2008, Poeppel et al. 2008, Abrams et al. 2009, Goswami 2011), it could be that synchronization to stimuli with less abrupt onsets (such as speech stimuli) depends more on slow-time-scale neural consistency. For stimuli with abrupt onsets, on the other hand, the auditory system may instead be able to rely on the robust onset responses produced by subcortical auditory regions.

We find that inter-trial phase-locking of the frequency following response and phase-locking of the lower-frequency cortical auditory evoked response are not correlated. This suggests that individual variations in phase coherence within the auditory system cannot be accounted for via a unitary mechanism. Instead, cortical and subcortical precision are decoupled, such that a given participant could have a precise low-frequency cortical response but an imprecise frequency-following response (or vice versa). Thus, a promising avenue for future research would be to investigate mechanisms that determine neural precision at fast and slow rates.

Although synchronization to an auditory beat is a skill that humans easily master, it is surprisingly rare in the animal kingdom, having been reported in only a handful of species to date (Patel et al. 2009, Hasegawa et al. 2011). Moreover, at present it has not been demonstrated in any other primate, despite at least one extensive attempt to teach the skill to rhesus monkeys (Zarco et al. 2009). The majority of the species that have been shown to synchronize, including the African Grey Parrot and the Sulphur-Crested Cockatoo, are capable of vocal learning (Patel et al. 2009, Schachner et al. 2009; although see Cook et al. 2013), suggesting that there may be overlap between the auditory-motor

connections which make vocal learning possible and those crucial for synchronization. Speculation to date regarding the necessary preconditions for synchronization has largely focused on connections between auditory cortical areas, pre-motor regions, and the basal ganglia (Patel et al. 2009, Merchant and Honing 2014, Patel and Iversen 2014). However, our finding that high-frequency auditory neural precision (> 100 Hz, corresponding to time scales of 10 ms and shorter) is linked to synchronization skill suggests that it may be fruitful to examine structural and functional interactions between subcortical auditory regions and motor areas as well. For example, strengthening of the direct connection between the auditory midbrain and cerebellum could enable the rapid, precise auditory-motor integration necessary for both vocal learning and synchronization. Conversely, a lack of a strong connection between subcortical auditory and motor areas in species that cannot perform vocal learning could explain the lack of a benefit for auditory versus visual stimuli for rhesus monkeys performing perceptual-motor synchronization tasks (Zarco et al. 2009, Kraus and White-Schwoch in press).

Our finding that auditory neural precision at a fast time scale enables rapid adaptation during synchronization adds to growing evidence that precision at fast time scales is crucial for auditory perception. Not only is neural precision linked to synchronization consistency (Tierney and Kraus 2013) and adaptation, but children with more consistent frequency-following responses also perform better at word-reading tasks (Hornickel and Kraus 2013). This suggests that a precise neural response enables extraction of a stable percept of the timing of sound and reliable categorization of speech sounds, while a less precise response interferes with these processes. Given that auditory neural precision appears to be important for the development of language skills and a gatekeeper for synchronization ability, it is an intriguing question whether synchronization training could enhance neural precision, potentially boosting language skills. Indirect support for this view is provided by studies of musical training, which have revealed that musical training tracks with enhanced consistency of neural responses to sound (Parbery-Clark et al. 2012, Skoe and Kraus 2013) and enhanced language skills (Moreno et al. 2009, Parbery-Clark et al. 2009, Kolinsky et al. 2009, Bhide et al. 2013, François et al. 2013, Zuk et al. 2013, Slater et al. 2014). Moreover, short-term auditory training has been shown to enhance both response consistency and language skills (Hornickel et al. 2012). Given that synchronization training via computer-based training programs would be cost-effective and efficient, examining the effects of synchronization training on neural function and language skill is a promising avenue for the future.
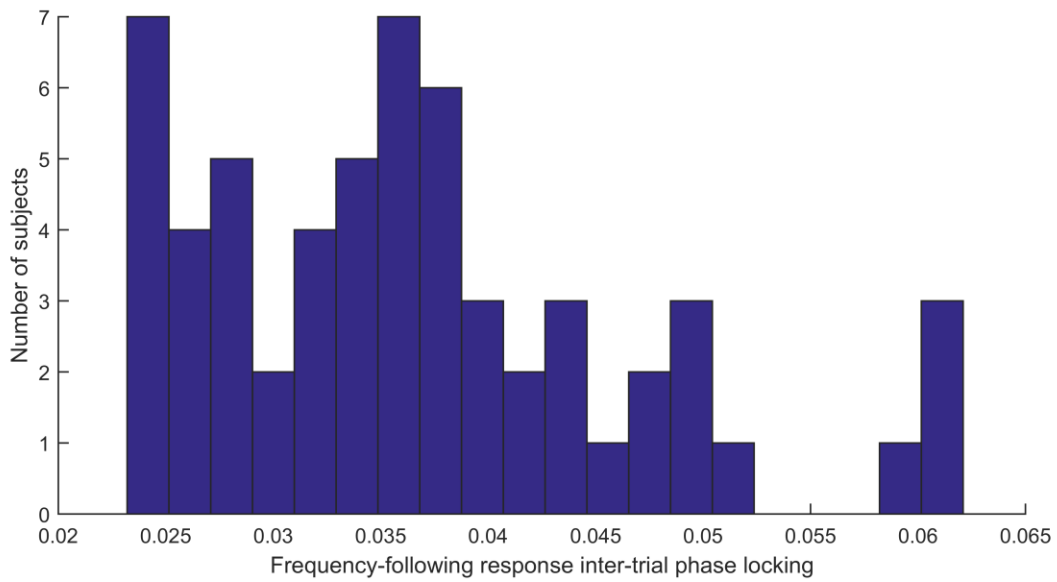
Figure 1. *Inter-trial phase locking across subjects*. Histogram of degree of inter-trial phase locking averaged across 20 Hz bins centered on the fundamental frequency and the first three harmonics (100, 200, 300, and 400 Hz).
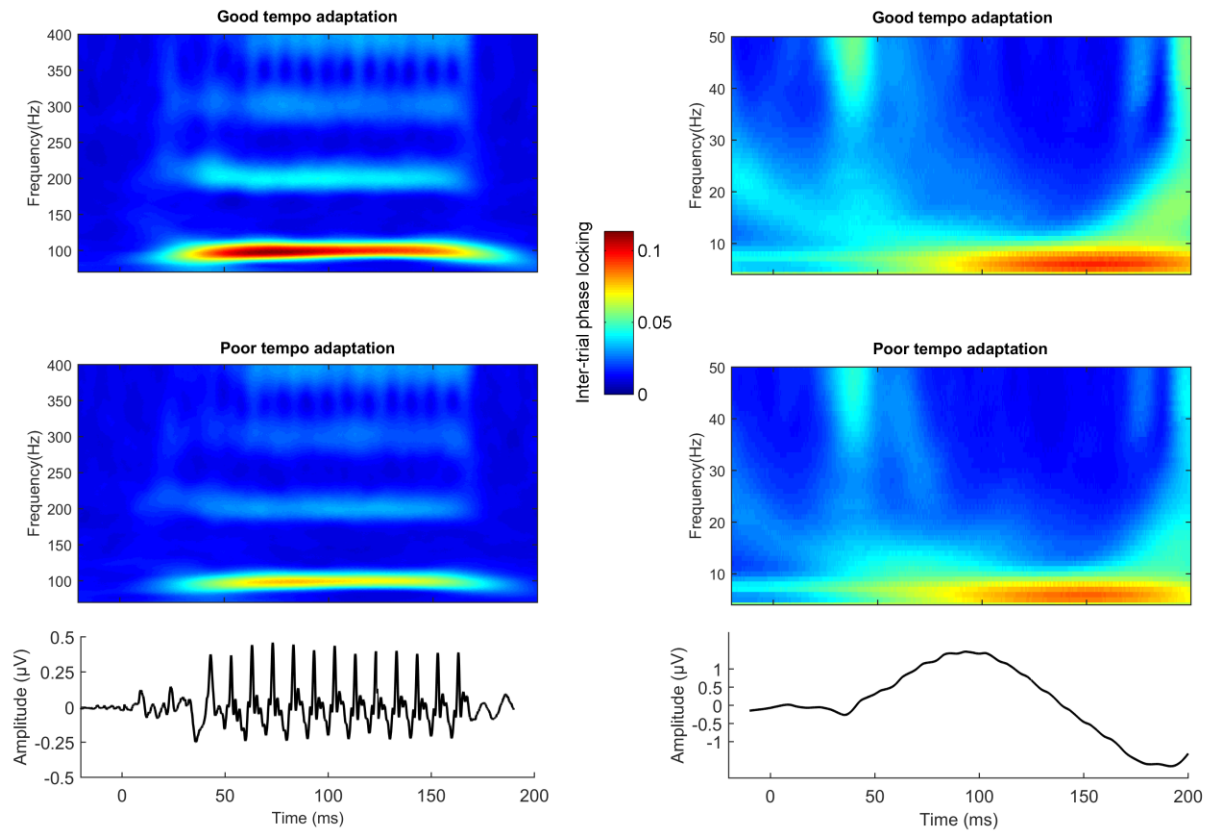
Figure 2. *Inter-trial phase locking and tempo adaptation performance.* (Upper Left) Inter-trial phase locking at 70 Hz and above in participants who performed better versus worse on a test of tempo adaptation (top versus bottom third performance). (Upper Right) Inter-trial phase locking at 50 Hz and below in participants who showed good versus poor tempo adaptation performance. (Bottom Left) Grand average response across all subjects high-pass filtered at 70 Hz. (Bottom Right) Grand average response low pass filtered at 50 Hz. Responses displayed in the bottom two panels are a composite of average responses to speech sounds presented in quiet and in background noise. Critical phase-locking values were 0.0223 for p = 0.05 as calculated using the method of Bonett and Wright (2000). Grand averages were calculated by taking the mean of the absolute values of the complex individual values.
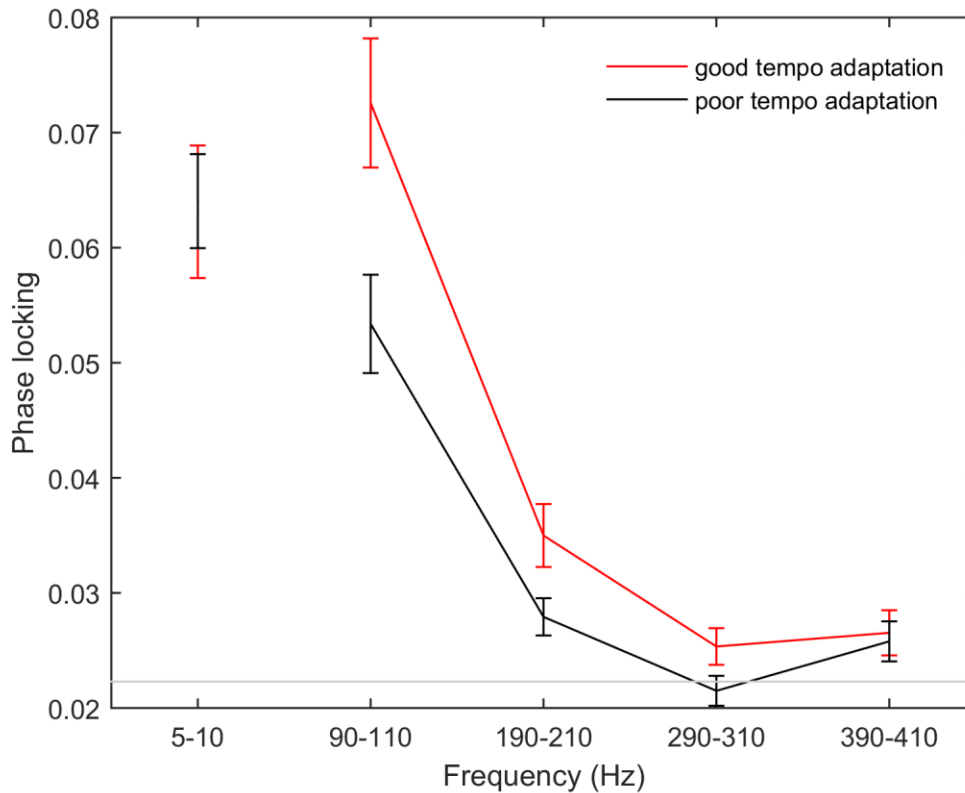

Figure 3. *Inter-trial phase locking and tempo adaptation performance.* Mean phase-locking values (composite score across both quiet and noise stimulus presentation conditions) for the participants who were better (red) versus worse (black) at quickly adapting to a change in metronome tempo (top versus bottom third performance). Behavioral performance was calculated as the absolute difference between the timing shift present in the stimulus and the timing shift produced by the participant. Error bars indicate one standard error of the mean. The light gray horizontal line indicates the critical value (0.0223) for significant phase locking at p = 0.05.
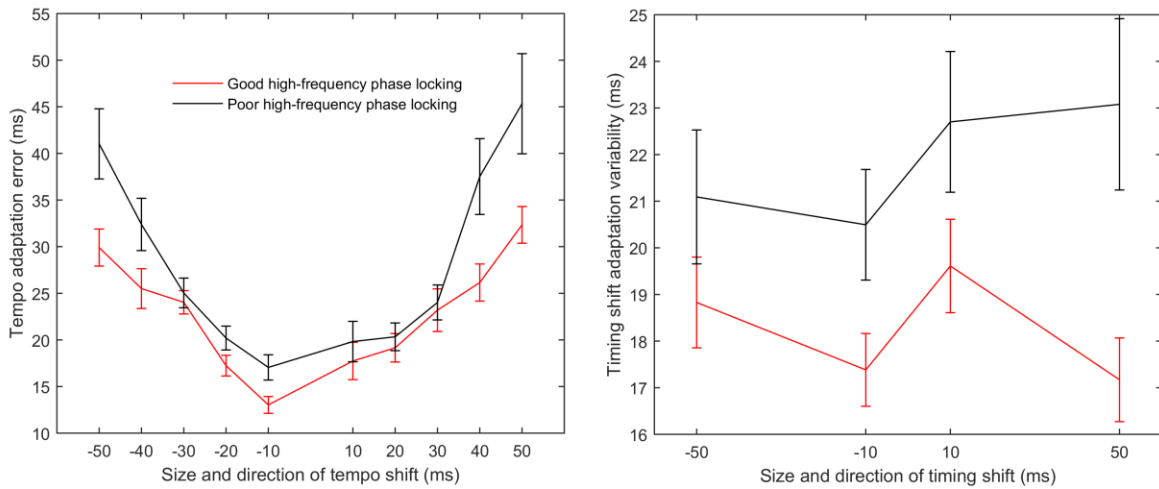
Figure 4. *Auditory-motor adaptation across conditions in participants with good versus poor high-frequency phase-locking.* Tempo (left) and timing shift (right) adaptation performance across all conditions in participants with good versus poor high-frequency phase-locking. Phase-locking was measured in 20 Hz bins centered on the fundamental frequency and the first three harmonics. Tempo adaptation performance was calculated as the absolute difference between the timing shift present in the stimulus and the timing shift produced by the participant. Timing adaptation performance was calculated as the standard deviation of the offset between the participant's movements and the drum sound onset for the six beats following the lengthened or shortened stimulus interval.
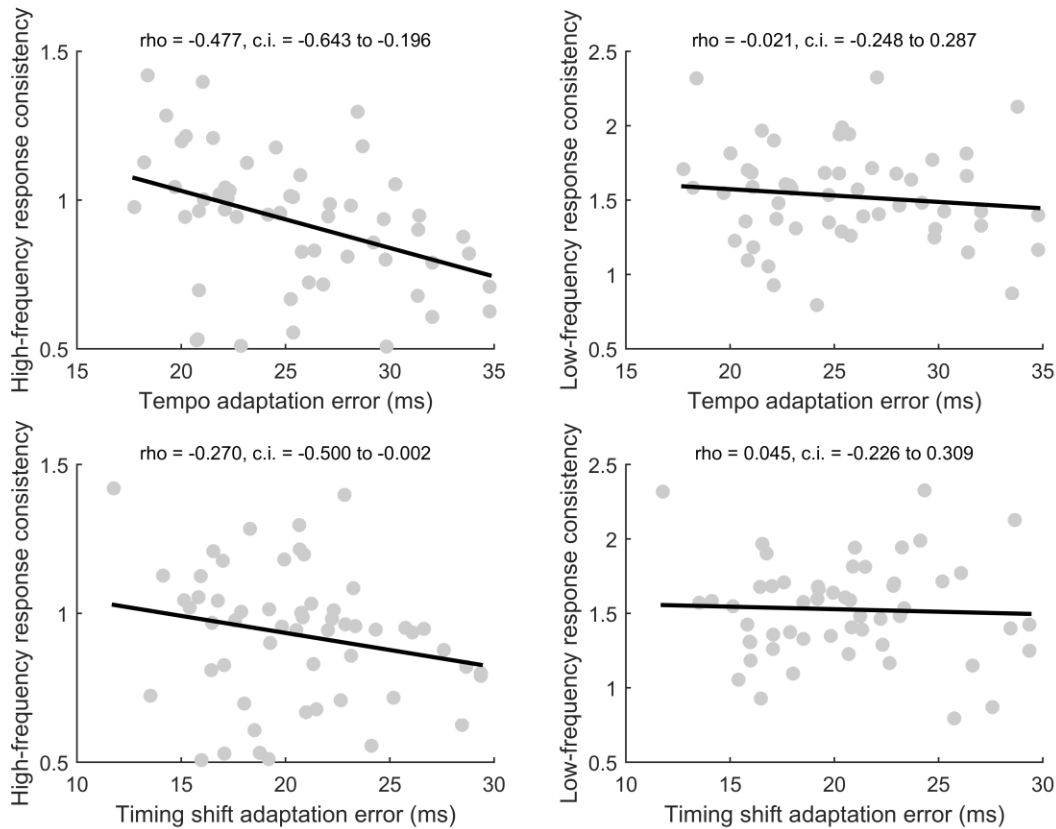
Figure 5. *Relationship between tempo and timing shift adaptation and response consistency.* (Top left) Scatterplot of tempo adaptation error versus frequency following response consistency (rho = -0.477, p < 0.001, confidence interval -0.643 to -0.196). (Top right) Scatterplot of tempo adaptation error versus low-frequency evoked response consistency (rho = -0.021, p = 0.881, confidence interval -0.248 to 0.287). These correlations were significantly different according to Meng's z-test (z = 2.801, p = 0.003). (Bottom left) Scatterplot of timing shift adaptation error versus frequency following response consistency (rho = -0.270, p = 0.044, confidence interval -0.500 to -0.002). (Bottom right) Scatterplot of timing shift adaptation error versus low-frequency evoked response consistency (rho = 0.045, p = 0.745, confidence interval -0.226 to 0.309). These correlations were significantly different according to Meng's z-test (z = 1.789, p = 0.037). Timing adaptation performance was calculated as the standard deviation of the offset between the participant's movements and the drum sound onset for the six beats following the lengthened or shortened stimulus interval.
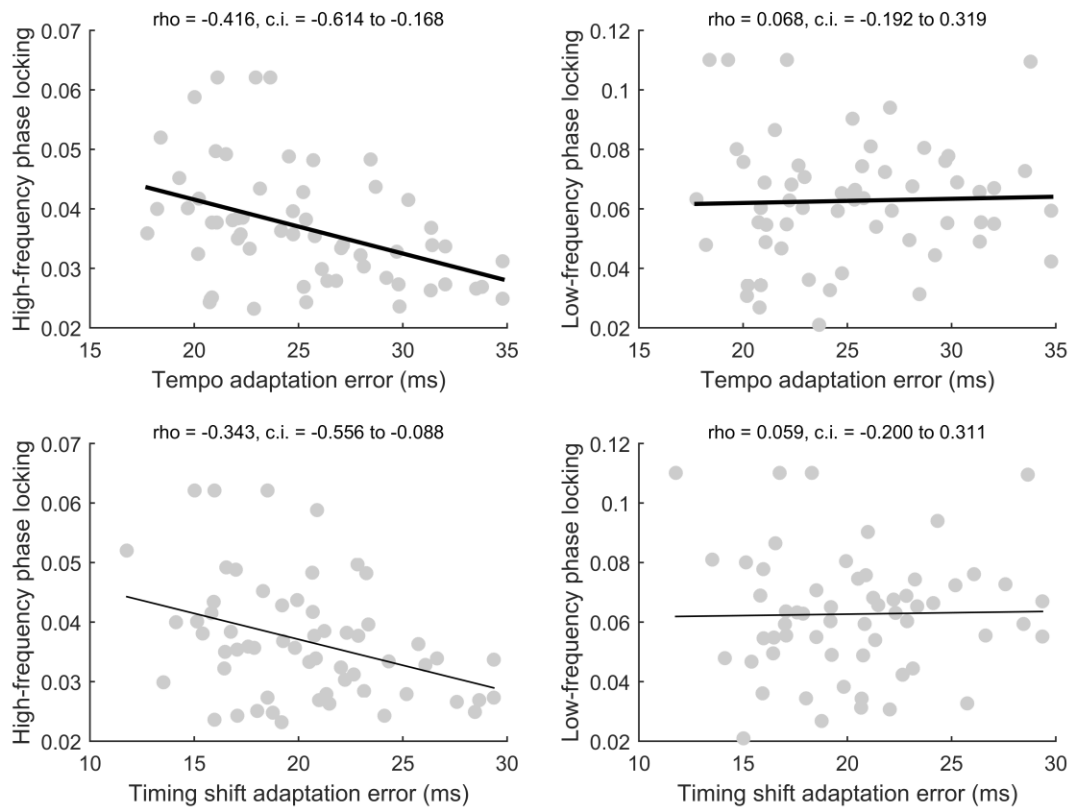
Figure 6. *Relationship between tempo and timing shift adaptation and inter-trial phase locking.* (Top left) Scatterplot of tempo adaptation error versus inter-trial phase locking measured in 20 Hz bins centered on the fundamental frequency and the first three harmonics (rho = -0.416, p = 0.001, confidence interval -0.614 to -0.168). (Top right) Scatterplot of tempo adaptation error versus inter-trial phase locking at 5-10 Hz (rho = 0.068, p = 0.608, confidence interval -0.192 to 0.319). These correlations were significantly different according to Meng's z-test (z = 2.735, p = 0.003). (Bottom left) Scatterplot of timing shift adaptation error versus high-frequency inter-trial phase locking (rho = -0.343, p = 0.008, confidence interval -0.556 to -0.088). (Bottom right) Scatterplot of timing shift adaptation error versus inter-trial phase locking at 5-10 Hz (rho = 0.059, p = 0.658, confidence interval -0.200 to 0.311). These correlations were significantly different according to Meng's z-test (z = 0.246, p = 0.012). Tempo adaptation performance was calculated as the absolute difference between the timing shift present in the stimulus and the timing shift produced by the participant.

Figure 7. *Relationship between tempo adaptation and timing shift adaptation performance.* Scatterplot of tempo adaptation versus timing shift adaptation error (rho = 0.477, p < 0.001, confidence interval 0.238 to 0.662). Phase-locking was measured in 20 Hz bins centered on the fundamental frequency and the first three harmonics. Tempo adaptation performance was calculated as the absolute difference between the timing shift present in the stimulus and the timing shift produced by the participant.
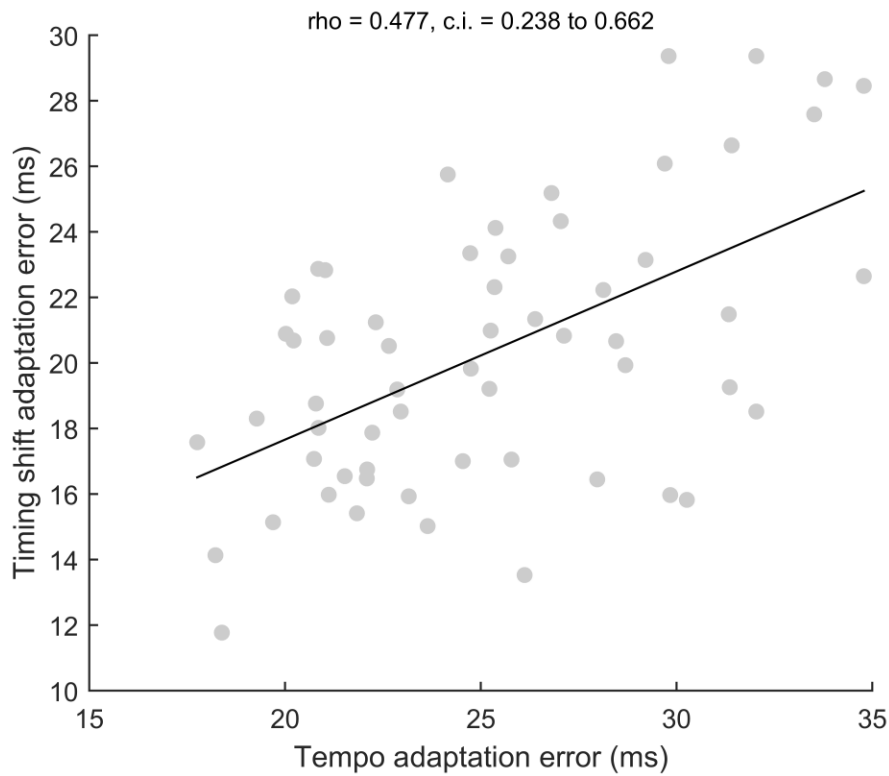
References

Abrams, D., Nicol, T., Zecker, S. & Kraus, N. (2008) Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *Journal of Neuroscience*, **28**, 3958-3965.

Abrams, D., Nicol, T., Zecker, S. & Kraus, N. (2009) Abnormal cortical processing of the syllable rate of speech in poor readers. *Journal of Neuroscience*, **29**, 7686-7693.

Aiken, S. & Picton, T. (2008) Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, **245**, 35-47.

Bajo, V., Nodal, F., Moore, D. & King, A. (2010) The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci*, **13**, 253-260.

Bartolo, R., Prado, L. & Merchant, H. (2014) Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *Journal of Neuroscience*, **34**, 3910-3923.

Bhide, A., Power, A. & Goswami, U. (2013) A rhythmic musical intervention for poor readers: a comparison of efficacy with a letter-based intervention. *Mind, Brain, and Education*, **7**, 113-123.

Bijsterbosch, J., Lee, K., Hunter, M., Tsoi, D., Lankappa, S., Wilkinson, I., Barker, A. & Woodruff, P. (2011) The role of the cerebellum in sub- and supraliminal error correction during sensorimotor synchronization: evidence from fMRI and TMS. *Journal of Cognitive Neuroscience*, **23**, 1100-1112.

Bonett, D., & Wright, T. (2000) Sample size requirements for estimating Pearson, Kendall and Spearman correlations. *Psychometrika*, **65**, 23-28.

Caclin, A., McAdams, S., Smith, B. & Winsberg, S. (2005) Acoustic correlates of timbre space dimensions: a confirmatory study using synthetic tones. *JASA*, **118**, 471-483.

Chandrasekaran, B. & Kraus, N. (2010) The scalp-recorded brainstem response to speech: neural origins. *Psychophysiology*, **47**, 236-246.

Chen, Y., Repp, B. & Patel, A. (2002) Spectral decomposition of variability in synchronization and continuation tapping: comparisons between auditory and visual pacing and feedback conditions. *Human Movement Science*, **21**, 515-532.

Chen, J., Zatorre, R. & Penhune, V. (2006) Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *NeuroImage*, **32**, 1771-1781.

Chen, J., Penhune, V. & Zatorre, R. (2008) Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, **20**, 226-239.

Coleman, J. & Clerici, W. (1987) Sources of projections to subdivisions of the inferior colliculus in the rat. *J Comp Neurol*, **262**, 215-226.

Cook, P., Rouse, A., Wilson, M. & Reichmuth, C. (2013) A California sea lion (zalophus californianus) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology*, **127**, 412.

François, C., Chobert, J., Besson, M. & Schön, D. (2013) Music training for the development of speech segmentation. *Cerebral Cortex*, **23**, 2038-2043.

Goswami, U. (2011) A temporal sampling framework for developmental dyslexia. *TICS*, **15**, 3-10.

Hasegawa, A., Okanoya, K., Hasegawa, T. & Seki, Y. (2011) Rhythmic synchronization tapping to an audio-visual metronome in budgerigars. *Scientific reports*, **1**, 120.

Hashikawa, T. (1983) The inferior colliculopontine neurons of the cat in relation to other collicular descending neurons. *J Comp Neurol*, **219**, 241-249.

Hornickel, J., Zecker, S., Bradlow, A. & Kraus, N. (2012) Assistive listening devices drive neuroplasticity in children with dyslexia. *Proceedings of the National Academy of Sciences*, **109**, 16731-16736.

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

Hove, M., Iversen, J., Zhang, A. & Repp, B. (2013) Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome. *Psychological Research*, **77**, 388-398.

Klemmer, E. (1967) Sequences of responses to signals engaged in time only. *Acta Psychologica*, **27**, 197-207.

Kolers, P. & Brewster, J. (1985) Rhythms and responses. Journal of Experimental Psychology: Human *Perception and performance*, **11**, 150-167.

Kolinsky, R., Cuvelier, H., Goetry, V., Peretz, I. & Morais, J (2009) Music training facilitates lexical stress processing. *Music Perception*, **26**, 235-246.

Kotz, S. & Schwartze, M. (2010) Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences*, **14**, 392-399.

Kraus, N. & White-Schwoch, T. (in press) Unraveling the biology of auditory learning: a cognitive-sensorimotor-reward framework. *Trends in Cognitive Sciences*.

Kudo, M. & Niimi, K. (1980) Ascending projections of the inferior colliculus in the cat: an autoradiographic study. *J Comp Neurol*, **191,** 545-556.

Large, E. (2008) Resonating to Musical Rhythm: Theory and Experiment. In Grondins S. (ed), *The Psychology of Time*. Emerald, West Yorkshire, pp. 189-232.

Lee, K., Egleston, P., Brown, W., Gregory, A., Barker, A. & Woodruff, P. (2007) The role of the cerebellum in subsecond time perception: evidence from repetitive transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, **19**, 147-157.

Liu, L., Palmer, A. & Wallace, M. (2006) Phase-locked responses to pure tones in the inferior colliculus. *J Neurophysiol*, **96**, 1926-1935.

Luo, H. & Poeppel, D. (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, **54**, 1001-1010.

Madison, G. & Merker, B. (2004) Human sensorimotor tracking of continuous subliminal deviations from isochrony. *Neuroscience Letters*, **370**, 69-73.

Malcolm, M., Lavine, A., Kenyon, G., Massie, C. & Thaut, M. (2008) Repetitive transcranial magnetic stimulation interrupts phase synchronization during rhythmic motor entrainment. *Neuroscience Letters*, **435**, 240-245.

Merchant, H., Pérez, O., Zarco, W. & Gámez, J. (2013) Interval tuning in the primate medial premotor cortex as a general timing mechanism. *Journal of Neuroscience*, **33**, 9082-9096.

Merchant, H. & Honing, H. (2014) Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, **7**, 1.

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M. & Fitch, T. (2015) Finding the beat: a neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B*, **370**, 20140093.

Molinari, M., Leggio, M., Filippini, V., Gioia, M., Cerasa, A. & Thaut, M. (2005) Sensorimotor transduction of time information is preserved in subjects with cerebellar damage. *Brain Research Bulletin*, **67**, 448-458.

Molinari, M., Leggio, M. & Thaut, M. (2007) The cerebellum and neural networks for rhythmic sensorimotor synchronization in the human brain. *The Cerebellum*, **6**, 18-23.

Moreno, S., Marques, C., Santos, A., Santos, M., Luis Castro, S. & Besson, M. (2009) Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cerebral Cortex*, **19**, 712-723.

Mower, G., Gibson, A. & Glickstein, M. (1979) Tectopontine pathway in the cat: laminar distribution of cells of origin and visual properties of target cells in dorsolateral pontine nucleus. *J Neurophysiol*, **42**, 1-15.

Nozaradan, S., Peretz, I., Missal, M. & Mouraux, A. (2011) Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience*, **31**, 10234-10240.

Nozaradan, S., Peretz, I. & Mouraux, A. (2012) Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, **32**, 17572-17581.

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

Parbery-Clark, A., Anderson, S., Hittner, E. & Kraus, N. (2012) Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Frontiers in Aging Neuroscience*, **4**, 1-12.

Patel, A., Iversen, J. & Chen, Y. (2005) The influence of metricality and modality on synchronization with a beat. *Exp Brain Res*, **163**, 226-238.

Patel A, Iversen J, Bregman M, Schulz I (2009) Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, **19**, 827-830.

Patel, A. & Iversen, J. (2014) The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, **8**, 57.

Peelle, J. & Davis, M. (2012) Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, **3**, 320.

Poeppel, D., Idsardi, W. & Wassenhove, V. (2008) Speech perception at the interface of neurobiology and linguistics. *Phil. Trans. R. Soc. B*, **363**, 1071-1086.

Pollok, B., Muller, K., Aschersleben, G., Schmitz, F., Schnitzler, A. & Prinz, W. (2003) Cortical activations associated with auditorily paced finger tapping. *Neuroreport*, **14**, 247-250.

Repp, B. (2000) Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychological Research*, **63**, 106-128.

Repp, B. (2003) Rate limits in sensorimotor synchronization with auditory and visual sequences: the synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, **35**, 355-370.

Saint Marie, R. (1996) Glutamatergic connections of the auditory midbrain: selective uptake and axonal transport of D-[^{3}H]aspartate. *J Comp Neurol*, **373**, 255-270.

Schachner, A., Brady, T., Pepperberg, I. & Hauser, M. (2009) Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, **19**, 831-836.

Schulze, K., Vargha-Khadem, F. & Mishkin, M. (2012) Test of a motor theory of long-term auditory memory. *PNAS*, **109**, 7121-7125.

Semjen, A., Vorberg, D. & Schulze, H. (1998) Getting synchronized with the metronome: comparisons between phase and period correction. *Psychol Res*, **61**, 44-55.

Skoe, E. & Kraus, N. (2013) Musical training heightens auditory brainstem function during sensitive periods in development. *Frontiers in Psychology*, **4**, 622.

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

Thaut, M., Miller, R. & Schauer, L. (1998) Multiple synchronization strategies in rhythmic sensorimotor tasks: phase vs period correction. *Biol. Cybern.*, **79**, 241-250.

Tierney, A., & Kraus, N. (2013) The ability to move to a beat is linked to the consistency of neural responses to sound. *Journal of Neuroscience*, **33**, 14981-14988.

Tierney, A. & Kraus, N. (2014) Neural entrainment to the rhythmic structure of music. *Journal of Cognitive Neuroscience*, **27**, 400-408.

Warren, J., Wise, R. & Warren, J. (2005) Sounds do-able: auditory-motor transformations and the posterior temporal plane. *TRENDS in Neurosciences*, **28**, 636-643.

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

Wolpert, D., Miall, R. & Kawato, M. (1998) Internal models in the cerebellum. *Trends in Cognitive Sciences*, **2**, 338-347.

Zarco, W., Merchant, H., Prado, L. & Mendez, J. (2009) Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *J Neurophysiol*, **102**, 3191-3202.

Zuk, J., Ozernov-Palchik, O., Kim, H., Lakshminarayanan, K., Gabrieli, J., Tallal, P. & Gaab, N. (2013) Enhanced syllable discrimination thresholds in musicians. *PLoS ONE*, **8**, e80546.