Repetition Suppression in Auditory–Motor Regions to Pitch and Temporal Structure in Music

Rachel M. Brown[1], Joyce L. Chen[2], Avrum Hollinger[1], Virginia B. Penhune[3],
Caroline Palmer[1], and Robert J. Zatorre[1]

Abstract

■ Music performance requires control of two sequential structures: the ordering of pitches and the temporal intervals between successive pitches. Whether pitch and temporal structures are processed as separate or integrated features remains unclear. A repetition suppression paradigm compared neural and behavioral correlates of mapping pitch sequences and temporal sequences to motor movements in music performance. Fourteen pianists listened to and performed novel melodies on an MR-compatible piano keyboard during fMRI scanning. The pitch or temporal patterns in the melodies either changed or repeated (remained the same) across consecutive trials. We expected decreased neural response to the patterns (pitch or temporal) that repeated across trials relative to patterns that changed. Pitch and temporal accuracy were high, and pitch accuracy improved when either pitch or temporal sequences repeated over trials.

Repetition of either pitch or temporal sequences was associated with linear BOLD decrease in frontal–parietal brain regions including dorsal and ventral premotor cortex, pre-SMA, and superior parietal cortex. Pitch sequence repetition (in contrast to temporal sequence repetition) was associated with linear BOLD decrease in the intraparietal sulcus (IPS) while pianists listened to melodies they were about to perform. Decreased BOLD response in IPS also predicted increase in pitch accuracy only when pitch sequences repeated. Thus, behavioral performance and neural response in sensorimotor mapping networks were sensitive to both pitch and temporal structure, suggesting that pitch and temporal structure are largely integrated in auditory–motor transformations. IPS may be involved in transforming pitch sequences into spatial coordinates for accurate piano performance. ■

INTRODUCTION

Music from all genres and cultures combine two major structural features, the ordering of tones (pitch sequences) and the temporal spacing between successive pitches (temporal or timing sequences; Palmer, 1997). The specific combination of pitch and timing sequences contribute to the perception of a single melody (Jones, 1987; Jones, Summerell, & Marshburn, 1987; Jones, Boltz, & Kidd, 1982), yet the two dimensions can also be perceived independently (Thompson, 1994; Peretz & Kolinsky, 1993; Palmer & Krumhansl, 1987). The productions of pitch and of timing sequences are also partially dissociable; musicians tend to make timing errors when auditory feedback is delayed and pitch errors when the serial ordering of pitches in the auditory feedback is altered (Pfordresher, 2003). The ways in which pitch and temporal structure in auditory sequences are mapped to the motor system in production remain poorly understood. This study compared the neural correlates of pitch and temporal production to illuminate the degree to which these dimensions are processed independently or together.

Behavioral evidence conflicts as to whether pitch and temporal structures are processed independently or in a unified way by listeners. Some evidence suggests that listeners are more sensitive to independent pitch or temporal features than to how the two features combine. Listeners' quality judgments of melodic segments were better predicted by how listeners rated the separate pitch or temporal content of the segments than by combined pitch and temporal ratings (Palmer & Krumhansl, 1987). Listeners were also better at detecting changes to melodic segments that introduced a novel pitch and/or duration to the musical segment than those that combined the same pitches and durations differently (Thompson, 1994). Other evidence suggests that listeners are sensitive to how pitch and temporal structures combine (Jones, 1987). Listeners were better at detecting pitch differences between two melodies whose temporal structures were more predictable (Jones et al., 1982). Listeners were also better at recognizing pitch patterns paired with the same rhythms (Jones et al., 1987) and discriminating between rhythms paired with different pitch sequences (Peretz & Kolinsky, 1993). These findings suggest that pitch and temporal structures may therefore be perceived as unified melodies. We investigate the way in which the two dimensions are processed in melody performance.

When musicians perform a melody, they must produce a series of actions while monitoring auditory feedback.

[1]McGill University, [2]Sunnybrook Research Institute, University of Toronto, [3]Concordia University

Auditory–motor integration appears to engage a network of brain regions including auditory and premotor cortex (PMC), SMA, and parietal regions (Baumann et al., 2007; Lahav, Saltzman, & Schlaug, 2007; Zatorre, Chen, & Penhune, 2007; Bangert et al., 2006; Hickok & Poeppel, 2004). Little is known about whether these regions respond differently to pitch and temporal structure in auditory sequences. Previous studies have compared the neural response to ordinal and temporal properties of well-learned motor sequences (Bengtsson, Ehrsson, Forssberg, & Ullén, 2004) or visually guided finger movement sequences, in which spatial information cued specific effector movements and temporal spacing between visual stimuli cued movement timing (Garraux et al., 2005; Sakai, Ramnani, & Passingham, 2002; Schubotz & von Cramon, 2001). Ordinal and temporal dimensions of motor or visual–motor sequences appear to be processed by partially distinct regions of a frontal–parietal network involved in sensorimotor mapping (Garraux et al., 2005; Bengtsson et al., 2004; Schubotz & von Cramon, 2001). In music performance, pitch structure provides ordinal information by signaling specific effector movement sequences on an instrument (e.g., keypresses on a piano), whereas temporal structure organizes movements in time without specifying effectors (Chen, Penhune, & Zatorre, 2008a; Zatorre et al., 2007; Bengtsson, Ehrsson, Forssberg, & Ullén, 2005). When trained pianists performed musical sequences from notation, different visuomotor networks were sensitive to the pitch versus the temporal structure of the music (Bengtsson & Ullén, 2006). Auditory–motor mapping for pitch and temporal structure may therefore engage different neural circuits. Alternatively, auditory–motor networks may respond to pitch and timing sequences in a melody as an integrated whole, in which case similar regions may be engaged by the two dimensions. We tested these alternatives using fMRI to compare how auditory–motor networks were engaged in transforming the pitch and temporal structure of auditory sequences into corresponding actions.

The first goal of this study was to determine which brain regions were involved in transforming the temporal structure of a melody into the temporal organization of corresponding movements. Several sensorimotor regions are sensitive to the temporal structure of visually guided motor sequences, including pre-SMA, PMC, BG, and cerebellum during specific attention to temporal information (Schubotz & von Cramon, 2001); the putamen during temporal sequence manipulation (Garraux et al., 2005); and inferior parietal, temporal, and ventral PMC (vPMC) as well as cerebellum when learning temporal sequences (Sakai et al., 2002). Pre-SMA, inferior frontal, and premotor regions were sensitive to the temporal structure of well-learned motor sequences (Bengtsson et al., 2004). Inferior frontal, inferior temporal, lateral occipital, and parietal regions were particularly sensitive to temporal information in musical notation during music performance (Bengtsson & Ullén, 2006). Similar motor regions including SMA, pre-SMA, and BG are sensitive to features of auditory temporal

structure during rhythm perception (Grahn & Brett, 2007), and cerebellum, premotor, parietal and dorso-lateral pFC are sensitive to temporal structure during short-term retention or synchronization with auditory rhythms (Chen, Penhune, & Zatorre, 2008b; Lewis, Wing, Pope, Praamstra, & Miall, 2004; Sakai et al., 1999). Among these regions, the PMC appears to be particularly sensitive to features of auditory temporal structure (Chen et al., 2008a, 2008b; Chen, Zatorre, & Penhune, 2006; Lewis et al., 2004). Response in dorsal PMC (dPMC) and auditory association cortex was specifically modulated by the saliency of metrical accents while participants tapped along with an isochronous rhythm (Chen et al., 2006). Response in dPMC was also functionally correlated with auditory cortical response while participants synchronized with rhythms of varying complexity (Chen et al., 2008b). PMC was also engaged by listening to rhythms (Chen et al., 2008a). Thus, PMC may interface between auditory temporal structure and movement timing. If the temporal structure of a melody is mapped to movement independently of pitch structure in music performance, then temporal structure may selectively engage dPMC.

The second goal of the study was to determine which brain regions were involved in transforming melodic pitch sequences into corresponding actions. Ordinal structure of visually guided motor sequences engaged SMA, primary motor, and somatosensory regions and cerebellum during specific attention to ordinal information (Schubotz & von Cramon, 2001), the cerebellum during ordinal sequence manipulation (Garraux et al., 2005), and superior parietal, medial-temporal, and occipital regions when learning an ordinal sequence (Sakai et al., 2002). Superior parietal cortex and dPMC as well as BG and cerebellum were sensitive to the ordinal structure of well-learned motor sequences (Bengtsson et al., 2004). During music production, superior temporal, medial occipital, and cingulate cortex were particularly engaged by ordinal information in musical notation (Bengtsson & Ullén, 2006). Studies emphasizing auditory pitch structure have implicated ventral frontal motor regions in mapping specific pitch sequences to specific action sequences. Nonmusicians trained to perform a piano melody showed greater response in vPMC and inferior frontal cortex when they listened to the learned pitch sequence as compared with a novel sequence with the same pitches (Lahav et al., 2007). Activity in vPMC was also related to how well nonmusicians learned to perform a novel melody but not random pitch sequences (Chen, Rae, & Watkins, 2012). Musicians engaged vPMC and inferior frontal regions while discriminating melodies based on pitch sequences or harmonies (Brown & Martinez, 2007). Superior parietal regions may also be sensitive to pitch structure in music; response in the intraparietal sulcus (IPS) predicted how well musicians and nonmusicians transformed pitch sequences into different musical keys (Foster & Zatorre, 2010a). Thus, whereas vPMC may match specific pitch sequences to specific action sequences, parietal regions may

transform pitch sequences into action-relevant coordinates. If pitch structure in a melody is mapped to movement independently of temporal structure in music performance, then pitch structure may selectively engage vPMC and parietal regions.

The third goal of the study was to directly compare the neural networks involved in pitch-motor mapping and timing-motor mapping. When both pitch and temporal structure are relevant to a musical task, they may be processed by similar frontal–parietal regions. Musicians engage dPMC, vPMC, pre-SMA, and parietal cortex when performing, reading, or listening to familiar harmonically and rhythmically complex musical sequences and when simultaneously imagining the corresponding movements or sounds on their instruments (Baumann et al., 2007; Meister et al., 2004). Musicians also engage dPMC and vPMC when synchronizing or planning to synchronize with auditory rhythms without pitch variation (Chen et al., 2008a). In a task that required musicians to generate novel pitch or temporal structure in melodies, generating either type of structure engaged overlapping regions of both dPMC and vPMC (Berkowitz & Ansari, 2008), suggesting that pitch and timing may be at least partially integrated during motor planning. Pitch and timing dimensions may therefore engage similar frontal–parietal regions.

In summary, pitch and temporal structure may be mapped to motor movements during performance as separate or integrated features. Thus, the transformation of pitch and temporal structure into motor movements may engage distinct or overlapping neural circuitry. In the current study, we measured BOLD signal while pianists performed an auditory–motor mapping task on an MR-compatible piano keyboard. Pianists listened to short melodies and subsequently played them back. The pitch or the timing sequences in the melodies either changed or remained constant (repeated) over consecutive trials. This type of design has been employed in previous studies to dissociate ordinal and temporal properties of visuo-motor sequences (Sakai et al., 2002). It is known that repeated events result in decreased activity in neurons that process those events, also called repetition suppression (Grill-Spector, Henson, & Martin, 2006). We therefore expected pitch or timing sequences that repeated over trials to cause decreased neural response in brain regions that process those features. If the motor system dissociates pitch and timing sequences, we predicted reduced response in vPMC or parietal regions when pitch structure repeated over trials and reduced response in dPMC when temporal structure repeated over trials. If the motor system integrates pitch and timing sequences, we predicted reduced response in similar premotor and parietal regions to either pitch or timing repetition. We also expected repetition to influence pitch and temporal performance accuracy. If pitch and timing sequences are processed separately, pitch repetition should improve pitch accuracy and timing repetition should improve temporal accuracy. If pitch and timing sequences are integrated, pitch and timing repetition should improve both pitch and temporal accuracy.

METHODS

Participants

Fourteen healthy right-handed pianists (10 men) who were 21.88 years old (18–29 years) with 14.47 years (10–24 years) of formal, private piano training and normal hearing participated in the study. Handedness was indicated via self-report. No participants possessed absolute (perfect) pitch (according to self-report and performance on an absolute pitch assessment). Participants' self-rated sight-reading abilities ranged from 2 to 5 on a scale of 1–5 ($M = 3.43$, $SE = 0.25$). All participants gave written informed consent before participating in the study, which was approved by the Montreal Neurological Research Ethics Review Board.

Equipment

The scanning task was performed on an MR-compatible electronic piano keyboard (Hollinger, 2008; Hollinger, Steele, Penhune, Zatorre, & Wanderley, 2007; Figure 1A) with 11 weighted keys, nine of which were used for the current study (E through C; Figure 1B). The keyboard was attached to an adjustable plastic frame that fastened to the scanning bed. The keyboard was free of ferromagnetic parts with all electronic components relegated to the control room outside the scanner environment. Acquisition of key presses was accomplished using fiber optic sensors, which are immune to the scanner's electromagnetic interference, and movable mirrors attached to each key. Sensors comprised emitter–receiver pairs of optical fibers and were connected to a custom optoelectronic acquisition and control board where light reflected by the movable mirrors on depressed keys was converted into electronic signals; these signals were then analyzed and converted into key triggers sent over USB to a laptop PC. Presentation software on the laptop PC used the key triggers to control the onset of audio files for pitches corresponding to each key on the keyboard. Thus, each keypress resulted in the corresponding pitch sound. The current study is one of the first to examine playback on an instrument that produces real-time auditory pitch feedback in the scanner. All sound was presented to participants binaurally through MR-compatible Etymotic insert earphones. Sounds were amplified and adjusted to a comfortable level for each participant.

Stimuli

Fifty-four novel melodies were presented during the course of the study: 14 practice melodies were presented during the prescan familiarization, and 40 test melodies were presented during the scanning task. All melodies were presented in a piano timbre. Each melody consisted of eight 500-msec tones and lasted between 2.5 and 3.5 sec from first to last note onset. All melodies consisted of a single melodic line for the right hand. During the prescan
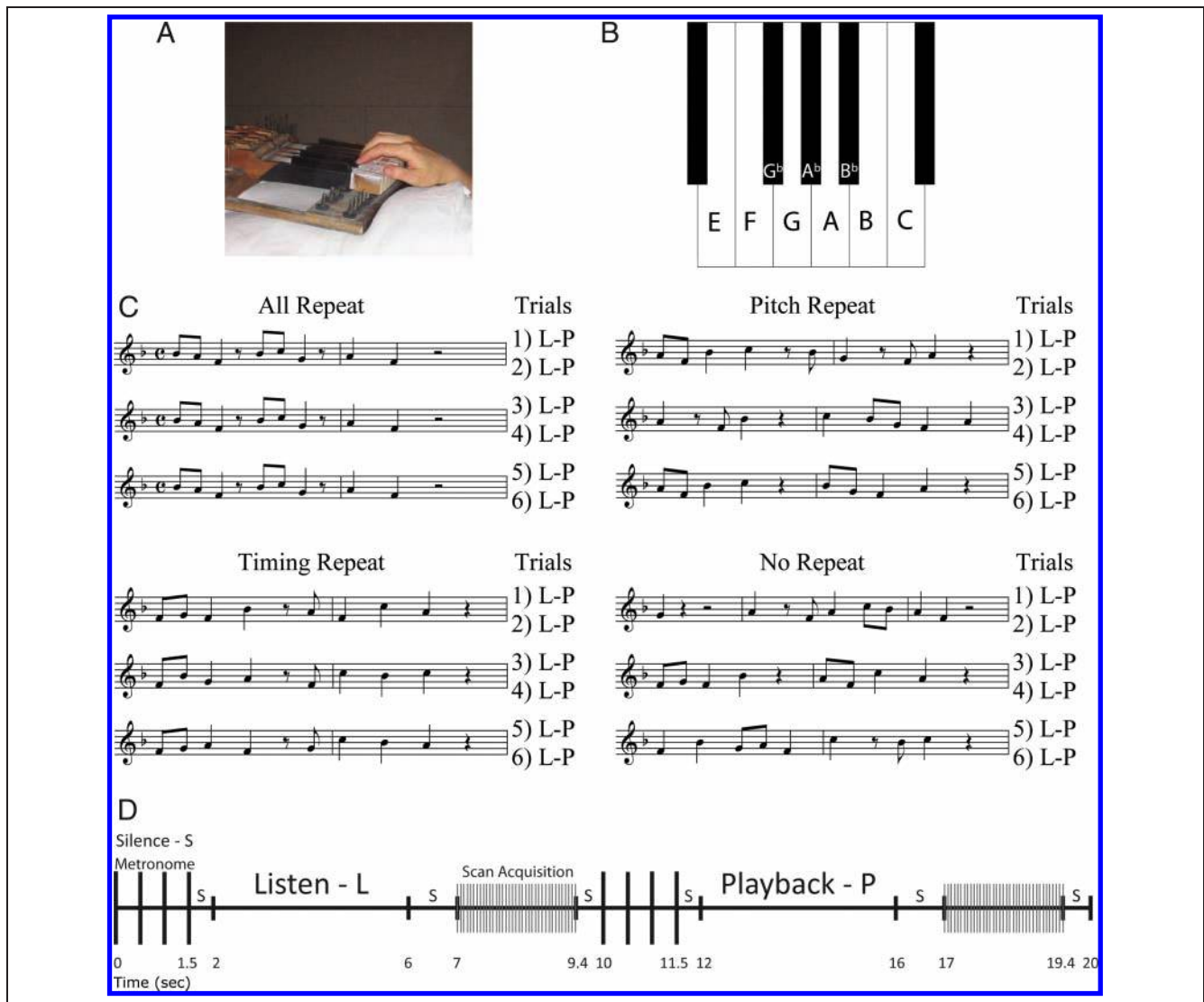
Figure 1. (A) fMRI-compatible keyboard. (B) Schematic of piano keys present on the keyboard and piano keys used for the current study (keys labeled with corresponding pitches). (C) Examples of scanning task blocks from each of the four conditions. In each condition, participants listened to (L) and subsequently played back (P) a melody six times, resulting in six Listen trials and six Playback trials per block (L-P × 6). In all conditions except the All Repeat condition, three different melodies were presented four times per block; thus, changes in pitch and/or timing sequences occurred every four trials (every two Listen trials and every two Playback trials). (D) Schematic of the sparse sampling paradigm: timing of events in Listen trials (L) and Playback trials (P) used in all conditions of the scanning task. Each trial began with four metronome beats (first 2 sec) followed by the onset of a melody (L) or a participant's performance of a melody (P). Listen or Playback occurred within a 4-sec window. This was followed by the scan acquisition (2.4 sec) sandwiched in between 1-sec and 0.6-sec silence buffers. Silence trials followed the same time course of events, with the exception that the 4-sec window between the metronome and scan acquisition consisted of silence. Key-cue trials were 10 sec each; in these trials, the metronome was omitted and verbal and musical cues were presented within the 6-sec time window before the scan acquisition.

and scanning tasks, each melody was preceded by four metronome beats (four 10-msec clicks presented in a drum timbre with an interonset interval of 500 msec). Tones and metronome clicks were generated in Cubase and output as WAV files, which comprised the stimuli and auditory feedback from the keyboard.

Melodies differed from one another according to the pitch sequence, the timing sequence (the sequence of interonset intervals or IOIs), or both. Fifty-four melodies were created by combining 40 unique pitch sequences with 39 unique timing sequences. Each pitch sequence contained tones from a unique set of five pitches; this allowed pianists to keep their hand in a single position on the keyboard when performing each melody, with one finger per pitch, thus minimizing gross hand or arm movements during performance. Each pitch sequence contained a total of eight pitches; there were no consecutive pitch repetitions. Each pitch sequence followed one of four musical keys: F major (14 sequences), E minor (14 sequences), C major (6 sequences), and A minor (6 sequences). Musical keys were not equally represented because the range of pitches available

on the keyboard constrained the number of possible pitch sequences in C major and A minor relative to F major and E minor (see Figure 1B). Each timing sequence was in 4/4 meter and contained a unique sequence of seven IOIs that were 1000, 750, 500, or 250 msec in duration (half, dotted-quarter, quarter, or eighth notes, respectively).

Task Design and Conditions

Pianists performed a listen–playback task in the scanner. Each trial consisted of either listening to a melody (Listen trial) or performing the melody that was heard on the previous trial by ear without notation (Playback trial). Listen and Playback trials were interleaved such that each Listen trial was followed by a single Playback trial and vice versa. Pianists always listened to and played back each melody twice to increase performance accuracy. Thus, each melody was heard and played back over four trials: two Listen and two Playback trials (Listen-Playback-Listen-Playback; Figure 1C). Trials were grouped into blocks that consisted of 12 trials: six Listen trials and six Playback trials, interleaved (Figure 1C). All Listen and Playback trials began with four metronome beats. Participants always heard their auditory feedback (all pitches and pitch onsets) during Playback trials. We used a listen–playback task rather than a sight-reading task to examine auditory–motor mapping processes without the influence of visual–motor or visual–auditory mapping processes. We therefore tested highly trained pianists who could perform melodies by ear with minimal error.

Task blocks varied according to whether the pitch and timing components of each melody changed or remained constant (repeated) over the course of a block (12 trials). This manipulation yielded four task conditions: (1) No Repeat (both the pitch and the timing sequence changed), (2) All Repeat (both the pitch and the timing sequence remained constant), (3) Pitch Repeat (only the timing sequence changed), and (4) Timing Repeat (only the pitch sequence changed; Figure 1C). In the No Repeat condition, both the pitch and timing sequence changed every four trials during the task block: Participants heard and played back a different pitch and timing sequence every four trials. Thus, the No Repeat condition contained three pitch sequence changes and three temporal sequence changes, and these changes happened simultaneously. In the All Repeat condition, both the pitch and timing sequence repeated over all 12 trials in a task block: Participants heard and played back the same melody during all trials. In the Pitch Repeat condition, only the timing sequence changed every four trials whereas the pitch sequence remained constant over all trials in the block: Participants heard and played back the same pitch sequence in all trials but a different timing sequence every four trials. In the Timing Repeat condition, only the pitch sequence changed every four trials whereas the timing sequence remained constant over all trials in the block: Participants heard and played back the same timing sequence in all trials but a different

pitch sequence every four trials. Thus, the Pitch Repeat and the Timing Repeat conditions both contained the same number of sequence repetitions and sequence changes: 12 pitch sequence repetitions and 3 temporal sequence changes in the Pitch Repeat condition, and 12 temporal sequence repetitions and 3 pitch sequence changes in the Timing Repeat condition.

The scanning task was divided into two runs. Each run consisted of eight task blocks (two per condition), eight Silence blocks, and two key-cue trials. Each Silence block lasted the equivalent of two task trials, and each key-cue trial lasted the equivalent of one task trial. Task and Silence blocks were interleaved, and each run always began with a Silence block. Each run contained 114 trials (96 task trials, 16 Silence trials, and 2 key-cue trials) and lasted 19 min (one run contained an extra 2 Silence trials at the end). Condition order across both runs was counterbalanced in a Latin-square fashion, and run order was counterbalanced across participants. The order of conditions was always the same within each run, thus maintaining the Latin-square condition order across the entire scan. To minimize hand movement during scanning, the entire task was blocked by musical key such that pianists only had to switch hand positions on the keyboard three times during the experiment. One run presented melodies in F major followed by C major, and the other run presented melodies in E minor followed by A minor. The first task block of each run as well as each musical key change within a run was preceded by a key-cue trial containing both a verbal auditory cue (the first author speaking the name of the key) and a musical auditory cue (a sequence of three pitches establishing the musical key: the first, third, and fifth scale degrees). The design was within subjects; the only between-subject factor was the order in which the two scanning runs were presented.

Sixteen unique pitch sequences and 16 unique timing sequences were presented during the scanning task. These pitch and timing sequences were combined to create 40 novel melodies that were presented during the scanning task. Pitch sequences were never combined with the same timing sequence more than twice, once for Listen and once for Playback, except during task blocks in the All Repeat condition. However, each individual pitch and timing sequence was presented the same number of times during the task: Each pitch sequence and each timing sequence was heard six times (Listen trials) and played back six times (Playback trials). Stimuli were presented this way to ensure equal exposure to each pitch and timing sequence (equal exposure to repeated sequences and nonrepeated sequences).

Procedure

Prescan

Participants were screened before scanning to make sure they could perform the listen–playback task with minimal error. Participants were trained to accurately execute each

of the four hand positions on the keyboard that corresponded with the four different musical keys. They then completed a short version of the scanning task using stimuli that were different from those presented during scanning but in the same musical keys. Participants completed the task on the same keyboard and computer used during scanning, and they completed the task while blindfolded to ensure that they could perform without visual input. Trial structure was identical to that of the scanning task, and scan acquisition noise was presented at the end of each trial to make sure that participants could overcome potential interference from the scanner noise between Listen and Playback trials. Participants were told to listen to each melody and play it back by ear on the following trial as accurately as possible. Participants were instructed to begin playing after the fourth metronome beat on playback trials. Participants who produced at least 85% of the pitches accurately during the playback trials were included in the study.

Scan

The keyboard was secured to the scanning bed at a comfortable arm's length for the participant. Padding was placed around participants' right (performing) arm and head to minimize movement. Participants were reminded of the hand position for each musical key and were blindfolded to minimize eye movements. Participants then performed the scanning task. All keystrokes and keystroke onsets produced during Playback trials were recorded on-line.

fMRI Acquisition

Scanning was performed on a 3-T Siemens Sonata Imager with a 32-channel head coil. A high resolution T1-weighted anatomical scan was first acquired for each participant (voxel size = $1 \times 1 \times 1$ mm^3, field of view = 224 mm^2). Two functional T2*-weighted gradient echo-planar runs were then acquired for each participant. One run contained 114 volumes and the other, 116 (due to two extra Silence trials at the end of the run). Each volume contained 40 whole-head interleaved slices (echo time = 30 msec, repetition time = 10,000 msec, voxel size = $3.5 \times 3.5 \times 3.5$ mm^3, matrix size = $64 \times 64 \times 40$, field of view = 224 mm^2); each slice was oriented perpendicular to the Sylvian fissure.

The two functional runs used a sparse-sampling paradigm, which minimizes the influence of the BOLD response due to scanner noise upon BOLD response to the task (Gaab, Gabrieli, & Glover, 2007; Belin, Zatorre, Hoge, Evans, & Pike, 1999). Volumes were acquired every 10 sec (repetition time = 10 sec) and took 2.4 sec to acquire. Stimulus presentation or performance took place within the 7.6 sec between scan acquisitions (Figure 1D). This paradigm takes advantage of the 4- to 6-sec delay in the hemodynamic response peak following a stimulus or event (Glover, 1999).

Behavioral Analyses

Performance on the Playback trials during scanning was assessed for pitch accuracy and temporal accuracy. Each measure was calculated separately for every Playback trial for each participant. Pitch accuracy was calculated as the percentage of correctly-produced pitches in each Playback trial. Omitted and substituted pitches were counted as errors. Temporal accuracy was calculated as the percentage of correctly produced IOIs in each Playback trial. Correct IOIs were defined as those which fell within a range defined by upper and lower limits set to halfway between the target IOI and neighboring target IOIs (126–374 for a target IOI of 250 msec, 376–624 for a target IOI of 500 msec, 626–874 for a target IOI of 750 msec, and 876–1124 for a target IOI of 1000 msec), similar to Drake and Palmer's (2000) coding of temporal errors.

To assess the change in performance accuracy across trials in each condition, change in performance accuracy from Trial 1 to successive trials was also examined. The first trial of every condition served as a baseline for subsequent trials because repetition or change manipulations occurred from Trial 2 onward. For each task block and performance accuracy measure, the first Playback trial value was subtracted from each subsequent Playback trial value (Trials 2–6) and divided by the first Playback trial value. This calculation yielded five percent change values for each performance accuracy measure (pitch accuracy and temporal accuracy) for each block of the scanning task.

fMRI Analyses

Functional MRI data were analyzed using the fMRI of the Brain Centre (FMRIB) Software Library (FSL, www.fmrib.ox.ac.uk/fsl; Smith et al., 2004). Functional images were preprocessed using FEAT (FMRIB's Expert Analysis Tool); images were motion corrected using MCFLIRT (Motion Correction FMRIB Linear Registration Tool; Jenkinson, Bannister, Brady, & Smith, 2002) and spatially smoothed using a Gaussian kernel of 8-mm FWHM. The first volume of each functional run and volumes pertaining to key-cue trials were discarded from analyses. A high-pass filter of 100 sec was used to remove low-frequency drift. Non-brain tissue was removed from functional and anatomical scans using BET (Brain Extraction Tool; Smith, 2002). Each participant's functional images were registered to their respective structural images using FLIRT (FMRIB's Linear Registration Tool; Jenkinson et al., 2002; Jenkinson & Smith, 2001) with 7 degrees of freedom. Each participant's structural images were registered to MNI-152 standard space using nonlinear registration (FNIRT: FMRIB's Nonlinear Registration Tool) with 12 degrees of freedom.

Statistical analysis was based on the general linear model. Statistical maps of activity corresponding to repetition suppression effects were computed using a linear contrast. Each parameter estimate represented a linear decrease in BOLD signal across the six Listen trials or

the six Playback trials in one of the four conditions (for a total of eight parameter estimates). For each condition, Trials 1–6 (Listen or Playback trials) were assigned the following contrast coefficients: 5, 3, 1, −1, −3, −5. These values represent an equal magnitude of decrease following each trial. All Silence trials were assigned values of 0. Thus, the z statistical maps for each parameter estimate represented voxels whose BOLD response over Listen or Playback trials showed a significant linear decrease, compared with silence, for one of the four conditions. This model was assumed to be the most conservative test of repetition suppression because it assumed a continuous decrease in response over all six trials. Because any changes in the pitch and/or timing sequence only occurred every other trial, both pitch and timing sequences repeated every two trials (Trials 1–2, 3–4, and 5–6) in each condition. A linear contrast across all six trials was therefore used to capture the repetition response of interest rather than response to repetition between every two trials. The above analyses were first performed at the subject level, separately for each run, and then averaged across runs for each participant using higher-level, fixed effects modeling in FEAT. Group averages were obtained by submitting each single-subject activation map into a stage 1 group analysis in FLAME (FMRIB's Local Analysis of Mixed Effects; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004). z Statistical images were thresholded using clusters determined by $z > 2.3$ and a corrected significance threshold of $p < .05$. Anatomical localization was determined using the Juelich histological atlas (Eickhoff et al., 2007), the Harvard–Oxford cortical and subcortical structural atlases, and the cerebellar atlas, which are part of the FSL software.

Repetition suppression response to pitch sequences and to timing sequences were compared at the group level in conjunction analyses between the Pitch Repeat and Timing Repeat conditions. Conjunction analyses were performed by taking the spatial intersection between above-threshold ($z > 2.3, p < .05$, corrected) statistical maps for the Pitch Repeat and Timing Repeat conditions (Nichols, Brett, Andersson, Wager, & Poline, 2005). Repetition suppression responses to pitch and timing were also contrasted in two subtractions: Pitch Repeat minus Timing Repeat (Pitch Repeat > Timing Repeat) and Timing Repeat minus Pitch Repeat (Timing Repeat > Pitch Repeat). Pitch Repeat and Timing Repeat conditions were also contrasted with the All Repeat condition to determine how responses to pitch or timing repetition were influenced by concurrent change in the other dimension. Each condition was also contrasted with the No Repeat condition to confirm that response was due to repetition. Each of the above subtractions was first performed at the subject level and then averaged across subjects. Each analysis described above was performed separately for Listen and Playback trials.

A post hoc ROI analysis was performed using Featquery in FSL to more closely examine the BOLD response to pitch and timing repetition and to examine the relationship between BOLD response and performance accuracy. For each subject, percent BOLD signal change at each Listen and Playback trial in each condition was averaged across a 7-mm-radius sphere centered on a peak voxel from contrasts of interest. To examine whether BOLD response during either listening or performance predicted behavioral performance, BOLD response in the ROIs at each Listen or Playback trial was correlated with pitch and temporal accuracy at each Playback trial in each condition.
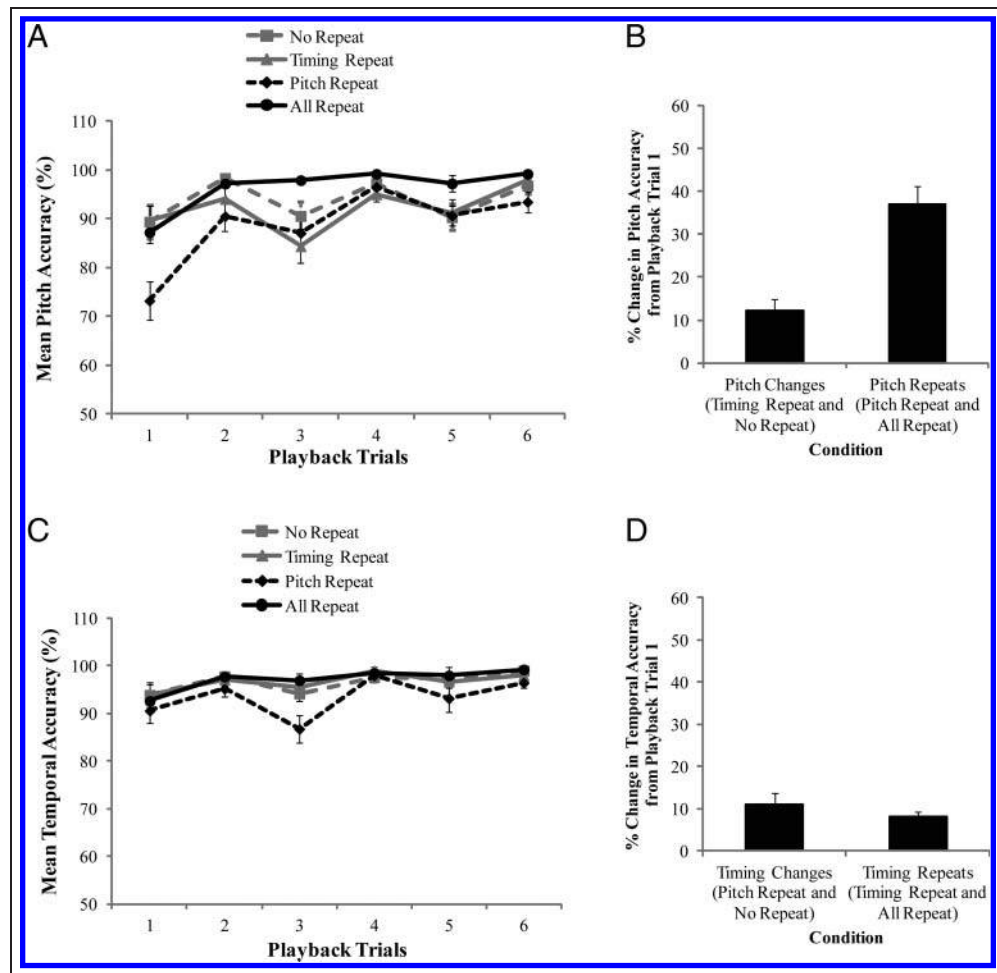
RESULTS

Behavioral Results

Pitch Accuracy

Mean pitch accuracy at each Playback trial in each condition is displayed in Figure 2A. Pitch accuracy was assessed in a 2 (Pitch Repetition: pitch sequence repeats or changes over trials) × 2 (Timing Repetition: timing sequence repeats or changes over trials) × 6 (Trial: Trials 1–6) repeated-measures ANOVA. The Pitch Repetition factor reflected a contrast between the mean of the All Repeat and Pitch Repeat conditions versus the mean of the Timing Repeat and No Repeat conditions; similarly, the Timing Repetition factor reflected a contrast between the mean of the All Repeat and Timing Repeat conditions versus the mean of the Pitch Repeat and No Repeat conditions (this is the case for all subsequent ANOVAs reported). An interaction between Pitch Repetition and Trial, $F(5, 65) = 9.45, p < .05$, indicated that pitch accuracy increased over trials when pitch repeated (Trials 2–6 > Trial 1; $HSD = 5.60, p < .05$) but not when pitch changed. An interaction between Timing Repetition and Trial, $F(5, 65) = 2.71, p < .05$, indicated that pitch accuracy also increased over trials when timing repeated (Trials 2, 4, 5, and 6 > Trial 1, Trials 4 and 6 > Trial 3; $HSD = 4.98, p < .05$) but not when timing changed. There was no three-way interaction. Thus, pitch accuracy increased when either pitch or timing sequences repeated over trials.

To examine how pitch or timing repetition influenced the magnitude of pitch accuracy improvement, percent change in pitch accuracy from Trial 1 was examined in a 2 (Pitch Repetition) × 2 (Timing Repetition) × 5 (Trial: Trials 2–6) repeated-measures ANOVA. Pitch accuracy was expected to improve more from Trial 1 to subsequent trials when pitch repeated versus when pitch changed over trials. This result was demonstrated by a main effect of Pitch Repetition, $F(1, 13) = 21.65, p < .05$: Percent change in pitch accuracy was greater when pitch repeated compared with when pitch changed over trials (Figure 2B). There was no main effect of Timing Repetition, indicating that the amount of pitch accuracy change over trials was not influenced by whether timing sequences repeated or changed over trials. There were

Figure 2. Pitch and temporal accuracy during Playback trials. (A) Mean pitch accuracy (percent correct) at each Playback trial in each of the four scanning task conditions. (B) Mean percent change in pitch accuracy from Playback Trial 1 to each subsequent Playback trial in conditions where pitch repeated (average of the Pitch Repeat and All Repeat conditions) compared with conditions where pitch changed (average of the Timing Repeat and No Repeat conditions). (C) Mean temporal accuracy (percent correct) at each Playback trial in each of the four scanning task conditions. (D) Mean percent change in temporal accuracy from Playback Trial 1 to each subsequent Playback trial in conditions where timing repeated (average of the Timing Repeat and All Repeat conditions) and conditions where timing changed (average of the Pitch Repeat and No Repeat conditions). Error bars represent standard error.

no two- or three-way interactions. Thus, the magnitude of pitch accuracy improvement was greater when pitch sequences repeated versus when pitch sequences changed over trials.

Temporal Accuracy

Mean temporal accuracy at each Playback trial in each condition is displayed in Figure 2C. Temporal accuracy was examined in a 2 (Pitch Repetition) × 2 (Timing Repetition) × 6 (trial) repeated-measures ANOVA. Temporal accuracy was expected to increase over trials when timing sequences repeated versus changed over trials. Temporal accuracy increased on average over trials, as indicated by a main effect of Trial, $F(5, 65) = 8.39, p < .05$ (Trials 2–6 > Trial 1, Trials 2, 4, 6 > Trial 3, $HSD = 3.43$, $p < .05$). Temporal accuracy was worst in the Pitch Repeat condition, as indicated by an interaction between Pitch Repetition and Timing Repetition, $F(1, 13) = 7.70, p < .05$ ($HSD = 2.79, p < .05$). There were no two-way interactions between Pitch Repetition and Trial or between Timing Repetition and Trial, and there was no three-way interaction. Thus, temporal accuracy did not benefit from either timing repetition or pitch repetition. Temporal accu-

racy was high overall ($M = 95.94\%$, $SE = 0.36$) and may have been near ceiling even at early trials. Participants' mean tempo was 512.70 msec ($SE = 0.86$) per quarter note, with a prescribed quarter note IOI of 500 msec; this suggests that participants adhered closely to the prescribed tempo during playback trials.

To examine how pitch or timing repetition influenced the magnitude of temporal accuracy improvement, percent change in temporal accuracy from Trial 1 was examined in a 2 (Pitch Repetition) × 2 (Timing Repetition) × 5 (Trial) repeated-measures ANOVA. Temporal accuracy was expected to improve more when timing repeated versus changed. Contrary to expectation, there was no main effect of Pitch or Timing Repetition. There was no interaction between Pitch Repetition and Trial. A main effect of Trial, $F(4, 52) = 7.49, p < .05$, and a two-way interaction between Timing Repetition and Trial, $F(4, 52) = 2.77, p < .05$, were driven by lowest accuracy improvement at Trial 3 than at other trials when timing did not repeat ($HSD = 5.72, p < .05$); temporal accuracy improvement did not differ across trials when timing repeated. Thus, the magnitude of temporal accuracy improvement was not sensitive to either pitch or timing repetition (Figure 2D).

Table 1. Brain Regions Showing Linear Response Decrease with Pitch and Timing Repetition

Brain Region	Listen Trials (x, y, z)	z	Playback Trials (x, y, z)	z
All Repeat Condition				
Pre-SMA	(−2, 6, 60)	4.02	(−10, 20, 38)	3.94
dPMC	(−20, 0, 54)	3.77	(−26, 2, 58)	3.1
			(24, −4, 52)	3.16
mid-PMC	(−42, −2, 44)	3.93		
vPMC/IFG	(−44, 2, 26)	4.38		
IFG			(−52, 8, 14)	3.38
			(50, 20, 8)	2.99
VLPFC			(−36, 26, −8)	4.01
			(38, 26, −8)	3.76
ACC			(8, 34, 12)	4.1
			(−6, 30, 20)	3.85
SPL	(−16, −62, 50)	4.16		
	(20, −62, 52)	4.39		
IPS	(−44, −36, 34)	4.58		
	(44, −36, 44)	4.1		
IPL	(−50, −34, 44)	4.36		
Insula	(−28, 26, 2)	3.12		
	(32, 26, 2)	3.26		
Caudate	(16, 20, 0)	3.57		
	(−14, 18, −2)	3.5		
Putamen	(−18, 14, −2)	3.61		

MNI coordinates of peak activations from the All Repeat condition and peak z values significant at $p < .05$, corrected. SPL = superior parietal lobule; IPL = inferior parietal lobule.

fMRI Results

Linear BOLD Decrease in the No Repeat Condition

As expected, no brain regions showed significant linear BOLD decrease in this control condition, either during Listen or Playback trials. No below-threshold activation was detected. This result suggests that the model of linear BOLD response decrease was appropriate for examining repetition suppression across the six Listen or Playback trials.

Linear BOLD Decrease in the All Repeat Condition

Brain regions whose BOLD response decreased linearly when both pitch and timing repeated over Listen trials included dPMC, pre-SMA, vPMC, mid-PMC, superior and inferior parietal cortex, insular cortex, and BG (Table 1). A

similar network of regions showed linear BOLD response decrease over Playback trials: dPMC, pre-SMA, and inferior frontal gyrus (IFG), as well as ACC and ventrolateral pFC (VLPFC; Table 1). Thus, repeated listening to or playback

Table 2. Brain Regions Showing Linear Response Decrease with Pitch Repetition

Brain Region	Listen Trials (x, y, z)	z	Playback Trials (x, y, z)	z
Pitch Repeat Condition				
Pre-SMA	(−2, 6, 60)	4.08	(−6, 24, 42)	3.41
dPMC	(−34, −2, 64)	4.47	(−24, 2, 70)	3.63
	(34, −2, 58)	4.38	(22, 12, 66)	3.87
mid-PMC	(−52, 0, 42)	3.72	(52, 2, 44)	2.85
	(52, 2, 46)	4.08		
vPMC/IFG	(−52, 10, 20)	4.25	(−58, 10, 36)	3.35
	(52, 10, 26)	4.01	(52, 8, 34)	2.59
IFG	(−48, 34, 14)	3.62	(−52, 30, 20)	3.56
	(54, 20, 24)	3.67		
MFG	(−32, 2, 64)	4.19	(36, 2, 64)	3.36
	(36, 2, 62)	4.08		
VLPFC	(−32, 24, −6)	3.30	(−36, 20, −12)	3.86
	(34, 26, −8)	3.10	(34, 26, 4)	3.65
Insula	(32, 24, 4)	4.07		
	(−32, 24, −2)	3.23		
SPL	(−24, −68, 54)	4.45	(−18, −68, 54)	3.42
	(12, −64, 64)	4.26	(16, −62, 62)	3.58
IPS	(36, −42, 42)	3.71		
	(−38, −38, 44)	3.78		
IPL	(−50, −36, 52)	3.92		
	(56, −38, 54)	3.56		
STG	(−60, −18, 4)	3.73		
	(60, −18, 2)	3.71		
Cerebellum				
Vermis VI	(2, −70, −14)	3.67	(−2, −82, −24)	3.04
Vermis VIIIa	(−2, −70, −42)	3.66		
Left VI	(−32, −40, −40)	3.62		
Left VIIb	(−28, −72, −58)	3.65		
Right VI	(28, −46, −36)	3.63	(10, −74, −20)	3.02
Right Crus I	(38, −72, −26)	3.73	(8, −82, −22)	2.93
Right VIIb			(12, −76, −44)	2.91

MNI coordinates of peak activations from the Pitch Repeat condition and peak z values significant at $p < .05$, corrected. SPL = superior parietal lobule; IPL = inferior parietal lobule.

of both pitch and timing sequences concurrently was accompanied by decreased BOLD response in frontal motor regions that primarily involved the PMC and pre-SMA.

Linear BOLD Decrease in the Pitch Repeat Condition

Regions whose BOLD response decreased linearly when only pitch repeated over Listen trials included dPMC, vPMC, mid-PMC, pre-SMA, IFG, middle frontal gyrus (MFG), VLPFC, superior parietal cortex, and the cerebellum, as well as the IPS, inferior parietal cortex, insular cortex, and the superior temporal gyrus (STG; Table 2). Similar regions showed linear BOLD response decrease over Playback trials: dPMC, vPMC, mid-PMC, pre-SMA, IFG, MFG, VLPFC, superior parietal cortex, and the cerebellum (Table 2, Figure 3A). Thus, repeated listening to or playback of pitch sequences was accompanied by decreased BOLD response in a frontal–parietal network, similar to the network that responded to concurrent pitch and timing repetition.

Linear BOLD Decrease in the Timing Repeat Condition

Regions whose BOLD response decreased linearly when only timing repeated over Listen trials included dPMC, pre-SMA, ACC, superior and inferior parietal cortex, and STG (Table 3). No brain regions showed above-threshold linear BOLD decrease over Playback trials. To examine whether this condition engaged a similar sensorimotor network as the other conditions,

z statistical maps were examined at a lower statistical threshold ($z > 2.3$, $p < .05$ uncorrected). Below-threshold linear BOLD decrease was detected in pre-SMA, dPMC, IFG, and VLPFC, as well as superior and inferior parietal cortex (Table 3, Figure 3B). Thus, timing repetition engaged similar frontal–parietal regions as pitch repetition or concurrent pitch and timing repetition, albeit less robustly.

Conjunction: Linear BOLD Decrease in the Pitch Repeat and Timing Repeat Conditions

To determine which brain regions responded similarly in the Pitch Repeat and Timing Repeat conditions, a conjunction analysis was performed between these conditions, separately for Listen and Playback trials. Regions showing linear response decrease in both conditions included dPMC, pre-SMA, STG, superior and inferior parietal cortex, and IPS during Listen trials (Figure 4A) and dPMC, pre-SMA, vPMC/IFG, superior and inferior parietal cortex, and VLPFC during Playback trials (Figure 4B). Thus, frontal motor regions and parietal regions responded similarly to pitch and timing sequence repetition during Listen and Playback trials, suggesting that a large part of the motor system responds to pitch and temporal structure as integrated features.

Subtraction: Contrast in Linear BOLD Decrease between the Pitch Repeat and Timing Repeat Conditions

Subtraction analyses were performed to determine how neural response decreases differed between the Pitch
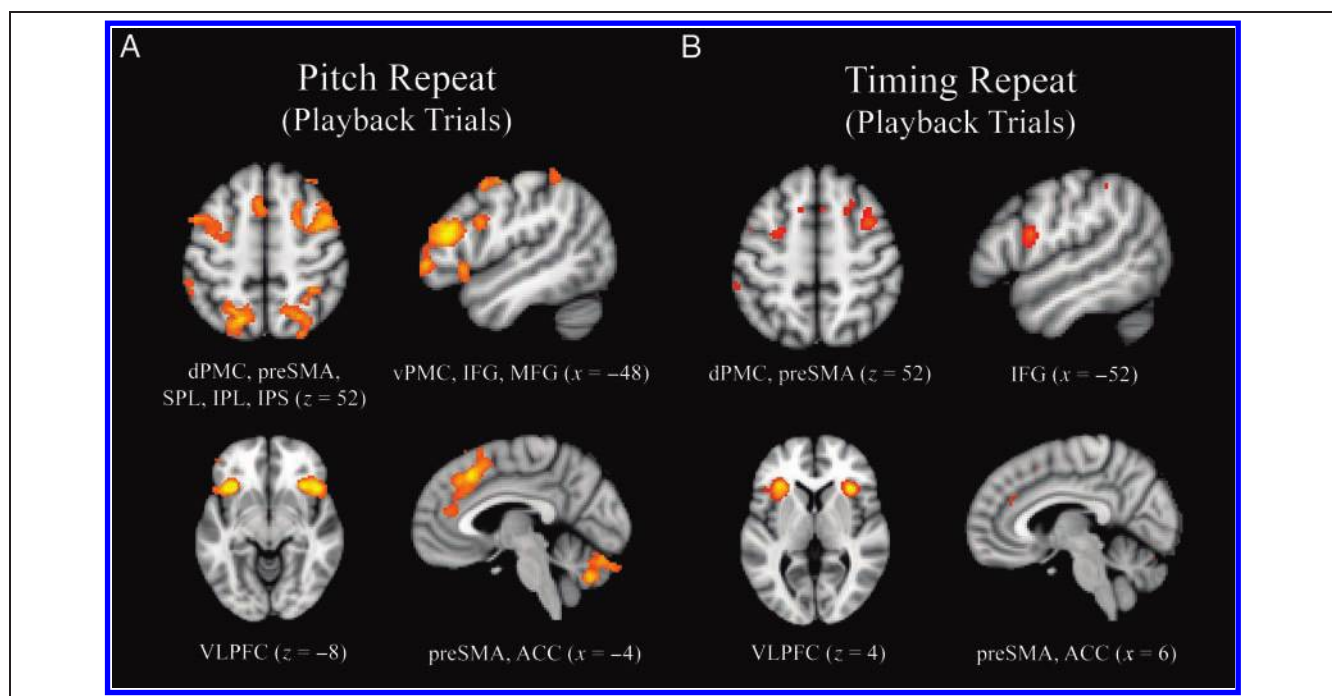
Figure 3. (A) z Statistical images, thresholded at $z > 2.3$ ($p < .05$, corrected), of brain regions showing linear BOLD response decrease during Playback trials in the Pitch Repeat condition. (B) z Statistical images (thresholded at $z > 2.3$, $p < .05$, uncorrected) of brain regions showing below-threshold linear BOLD response decrease during Playback trials in the Timing Repeat condition.

Table 3. Brain Regions Showing Linear Response Decrease with Timing Repetition

Brain Region	Listen Trials		Playback Trials[a]	
	(x, y, z)	z	(x, y, z)	z
Timing Repeat Condition				
Pre-SMA	(0, 10, 54)	3.26	(−10, 14, 54)	2.33
	(−4, 2, 66)	2.99		
dPMC	(−10, 10, 66)	3.19	(−16, 8, 60)	3.08
	(18, 8, 60)	2.81	(36, 8, 50)	2.92
IFG			(−50, 6, 20)	2.98
VLPFC			(−32, 26, 4)	3.36
			(34, 22, 6)	3.53
Frontal Pole			(−32, 48, 24)	2.89
			(36, 46, 28)	3.01
ACC	(10, 24, 28)	3.03		
SPL	(−6, −60, 68)	3.48	(−10, −56, 70)	2.97
			(26, −56, 58)	2.82
IPL	(−44, −32, 42)	3.05	(−60, −32, 40)	2.78
IPS	(−44, −34, 38)	2.68	(−34, −46, 42)	2.67
STG	(−58, −18, −4)	3.62		
	(66, −16, 4)	3.57		

MNI coordinates of peak activations from the Timing Repeat condition and peak z values significant at $p < .05$, corrected. SPL = superior parietal lobule; IPL = inferior parietal lobule.

[a]For Playback trials, peak z values are thresholded at $z > 2.3$ and are significant at $p < .05$, uncorrected.

Repeat and the Timing Repeat conditions. Subtraction of the Pitch Repeat condition from the Timing Repeat condition revealed no significant differences. Subtraction of the Timing Repeat condition from the Pitch Repeat condition revealed significantly greater linear decrease in bilateral superior and inferior parietal cortex, including bilateral IPS, in the Pitch Repeat condition (Table 4, Figure 5A). An additional contrast between the All Repeat condition and the Timing Repeat condition (All Repeat > Timing Repeat) also revealed significant linear BOLD response decrease in IPS. Subtraction of the All Repeat condition from the Pitch Repeat condition revealed no significant differences in IPS response. Together, these contrasts suggest that IPS response was sensitive to pitch repetition regardless of whether timing changed or not.

ROI Analysis (IPS)

To illustrate BOLD response decrease in IPS to pitch and timing repetition, percent BOLD signal change was extracted from ROIs centered around left and right peak voxels in IPS from the Pitch Repeat > Timing Repeat contrast (left: −40, −58, 46; right: 38, −40, 50). Percent BOLD signal change was extracted from the left and right ROI for each Listen and Playback trial in each condition; values from left and right ROIs were then averaged, as results were similar for either ROI separately. Mean percent change in BOLD response across right and left IPS in each condition is plotted over Listen and Playback trials in Figure 5B. This graph further illustrates the results of the subtraction analyses above: BOLD response in IPS decreased over Listen trials only when pitch repeated.

Because IPS response was sensitive to pitch repetition but not timing repetition, we examined whether IPS response during either listening or performing influenced participants' ability to perform pitch sequences correctly. Each participant's mean pitch accuracy score at each of the six Playback trials was correlated with each participant's mean percent BOLD signal change in IPS at each of the six Playback or Listen trials, separately in each condition. Both Pearson's and Spearman's rank correlations were conducted due to the nonnormality of the pitch accuracy score distribution. Pitch accuracy correlated negatively with IPS response during Listen trials (Pearson's $r = −.36$, Spearman's $r = −.32$, $ps < .05$; Figure 5C) and during Playback trials (Pearson's $r = −.25$, Spearman's $r = −.32$, $ps < .05$; Figure 5D), only in the Pitch Repeat condition. As BOLD signal in IPS decreased with pitch repetition over trials, pitch accuracy increased. IPS response did not correlate with pitch accuracy in any other condition. Thus, IPS response during both planning (listening) and execution (performance) influenced participants' ability to accurately produce pitch sequences across consecutive repetition trials, when pitch sequences repeated across trials.

DISCUSSION

The aim of the current study was to directly compare how two basic levels of musical sequence structure, pitch structure and temporal structure, are transformed into corresponding actions. We used a repetition suppression paradigm to identify brain regions sensitive to the two features. Behaviorally, both pitch and temporal accuracy improved across trials, validating the use of the repetition suppression paradigm. Improvement in pitch accuracy was facilitated by either pitch or temporal repetition but more so by pitch repetition. Repetition of pitch or temporal sequences corresponded to linear BOLD decrease in dPMC and pre-SMA, as well as vPMC, parietal cortex, and VLPFC. For Listen trials only, pitch sequence repetition was associated with linear BOLD decrease in IPS. The BOLD response decrease in IPS during Listen and Playback trials predicted pitch accuracy improvement during pitch but not temporal sequence repetition. Overall, the results demonstrate that frontal–parietal networks are similarly sensitive to both pitch and temporal structure but that parietal regions are more responsive to pitch structure. Thus,
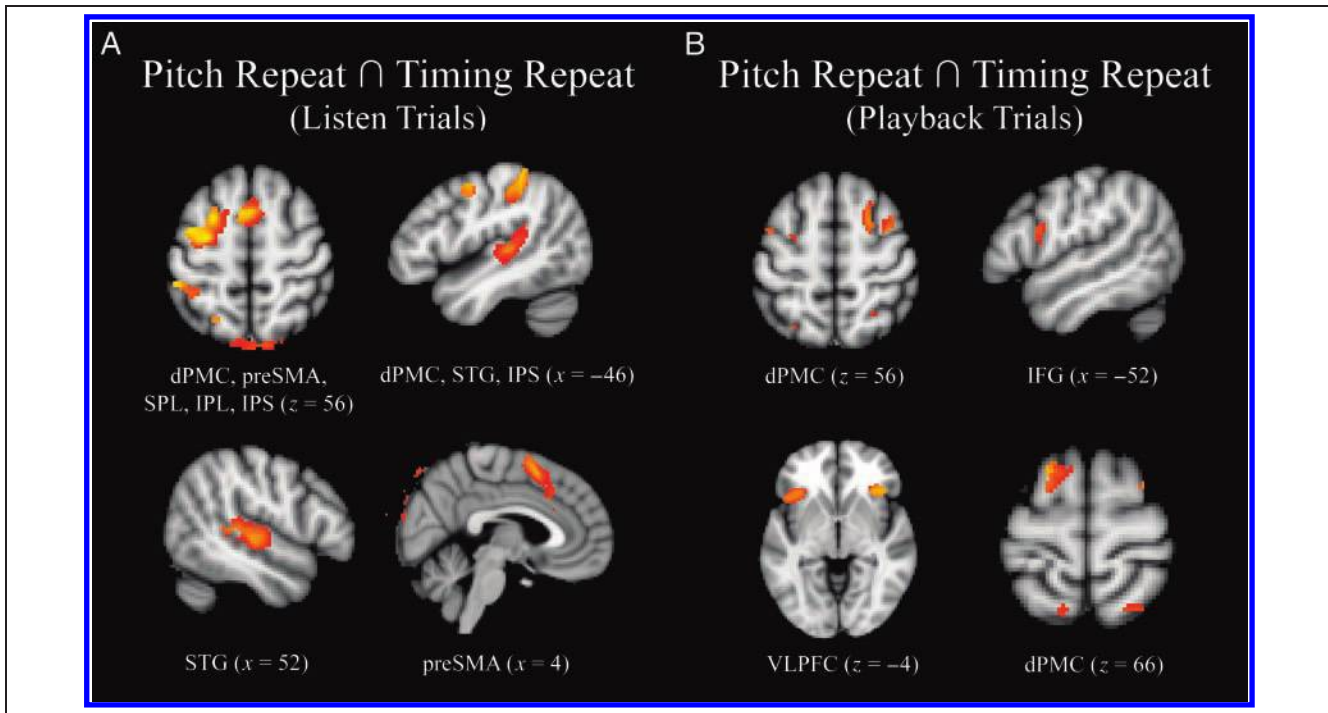
Figure 4. (A) Conjunction between *z* statistical maps of linear BOLD response decrease in the Pitch Repeat and Timing Repeat conditions for Listen trials. (B) Conjunction between above-threshold *z* statistical maps of linear BOLD response decrease in the Pitch Repeat condition during Playback trials (*z* > 2.3, *p* < .05, corrected) and below-threshold *z* statistical maps of linear BOLD decrease in the Timing Repeat condition during Playback trials (*z* > 2.3, *p* < .05, uncorrected).

these findings suggest that pitch and temporal structure are largely integrated in auditory–motor transformations.

Pitch and Temporal Accuracy

Pianists' performance accuracy during playback trials was similarly influenced by pitch and temporal structure; repetition of either dimension improved performance. The results suggest that pitch and temporal processing interacted, because timing repetition influenced pianists'

Table 4. Brain Regions Showing Greater Linear Response Decrease with Pitch Repetition than Timing Repetition

	Listen Trials	
Brain Region	(x, y, z)	z
Pitch Repeat > Timing Repeat		
SPL	(−24, −68, 54)	3.38
	(18, −62, 54)	3.4
IPL	(56, −42, 54)	3.52
	(−64, −36, 38)	2.90
IPS	(−40, −58, 46)	3.1
	(38, −40, 50)	2.99

MNI coordinates of peak activations and peak *z* values significant at *p* < .05, corrected. SPL = superior parietal lobule; IPL = inferior parietal lobule.

ability to play the correct sequence of pitches. Temporal accuracy was high overall and did not benefit from pitch or temporal repetition. Only pitch showed a dimension-specific effect in which pitch accuracy improved more with pitch sequence repetition than with temporal sequence repetition; this suggests some separability of pitch and temporal processing.
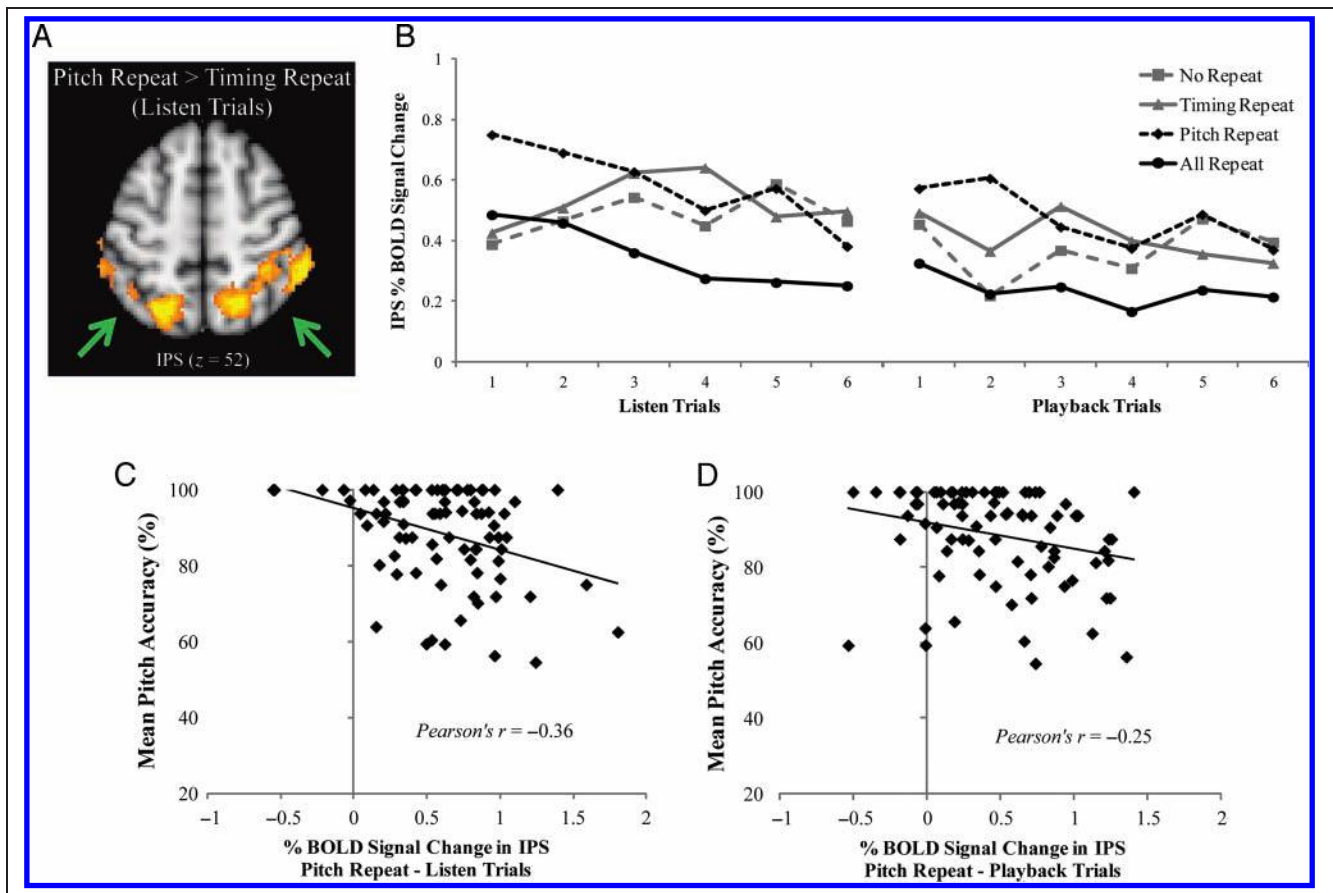
BOLD Response to Pitch and Timing

Frontal motor regions, including dPMC and pre-SMA, were similarly responsive to both pitch and timing repetition during listening and performance, suggesting that these regions process the two dimensions together. Previous studies of auditory rhythm reproduction (without pitch variation) have implicated both of these regions in the temporal organization of movement (Chen et al., 2006, 2008a, 2008b; Grahn & Brett, 2007; Lewis et al., 2004; Sakai et al., 1999) as well as dPMC in both pitch and rhythm production (Berkowitz & Ansari, 2008). Thus, dPMC and pre-SMA may have a generalized role in integrating multiple sensory cues that are relevant to a single unified action (Hoshi & Tanji, 2007), as well as in selecting from multiple stimulus-cued actions or response options (Cisek & Kalaska, 2005; Grafton, Fagg, & Arbib, 1998). The current results are also consistent with the proposed role of pre-SMA in sequential organization of actions (Sakai, Hikosaka, & Nakamura, 2004; Janata & Grafton, 2003) or conflict resolution between multiple motor plans (Nachev, Wydell,

Figure 5. (A) z Statistical map (thresholded at $z > 2.3$, $p < .05$, corrected) of brain regions showing greater linear BOLD response decrease in the Pitch Repeat than in the Timing Repeat condition (Pitch Repeat > Timing Repeat) during Listen trials. (B) Mean percent BOLD signal change in IPS at each Listen and Playback trial in each condition. (C) Pearson correlation between mean percent BOLD signal change in IPS at each Listen trial and mean pitch accuracy for each participant ($n = 84$) in the Pitch Repeat condition. (D) Pearson correlation between mean percent BOLD signal change in IPS at each Playback trial and mean pitch accuracy for each participant ($n = 84$) in the Pitch Repeat condition.

O'Neill, Husain, & Kennard, 2007); coding for both pitch-related and temporally related motor plans is consistent with either of these functions. In addition to dPMC, vPMC was also sensitive to both pitch and temporal structure, although less strongly to temporal structure. This result further suggests a role of vPMC in processing temporal cues for movement (Chen et al., 2006, 2008a, 2008b), not just pitch cues (Brown & Martinez, 2007; Lahav et al., 2007). Overall, the similarity in frontal motor response to pitch and timing suggests that these dimensions are mainly processed together when musicians are using auditory information to produce movement.

Response decrease to pitch and temporal structure was also similar in superior and inferior parietal regions and VLPFC. Parietal cortex forms part of the dorsal "action" processing stream (Rauschecker & Scott, 2009; Hickok & Poeppel, 2004; Goodale & Milner, 1992) and was likely involved in transforming the pitch and temporal dimensions of sound into motor-relevant coordinates. Response decrease in VLPFC to pitch or temporal repetition may have reflected decreasing memory retrieval demands during the task. VLPFC is thought to be engaged in active

memory retrieval requiring top–down control or selection among options (Kostopoulos & Petrides, 2003; Petrides, Alivisatos, & Evans, 1995). Such a retrieval process may have been more strongly engaged during early task trials when memory demands are greatest.

BOLD Response Decrease in IPS

Bilateral regions of IPS showed significant response decrease during pitch repetition compared with timing repetition. Response decrease in this region also predicted increase in pitch accuracy over trials during pitch repetition only. This region has been associated with spatial processing (Husain & Nachev, 2007) and mental rotation of visual objects (Zacks, 2008; Jordan, Heinze, Lutz, Kanowski, & Jäncke, 2001). However, IPS may play a more general role in reorganizing or transforming multimodal information (Foster & Zatorre, 2010a; Cusack, 2005; Grefkes, Weiss, Zilles, & Fink, 2002). This region is also engaged in auditory sequence transformations such as imagining temporally reversed melodies or mentally transposing melodies into different musical keys (Foster

& Zatorre, 2010a, 2010b; Zatorre, Halpern, & Bouffard, 2010). IPS receives inputs from multiple sensory regions (Frey, Campbell, Pike, & Petrides, 2008) and has been engaged in cross-modal object recognition (Grefkes et al., 2002). IPS may therefore transform structures into different, cross-modal coordinate systems while preserving the relationship among elements in the structure (Foster & Zatorre, 2010a; Grefkes, Ritzl, Zilles, & Fink, 2004). In the current study, this region may have been involved in transforming pitch sequences into spatial coordinates on the keyboard. Pianists may have imagined musical notation as they performed the task, which may have also engaged the IPS (Meister et al., 2004), although the crucial coordinate transformation in the current task was that of sound to spatial coordinates. Parietal response to pitch repetition was only greater during Listen trials, suggesting that the transformation from sound to space may have taken place mainly while pianists were planning their upcoming movements. Nonetheless, IPS response decrease during both listening and performance predicted pitch accuracy improvement suggesting that it is involved in both planning and performance.

Overall, the results suggest that pitch and temporal structures are largely integrated in auditory–motor transformations in music performance, which is consistent with behavioral evidence for pitch-timing integration in melody perception and memory (Jones, 1987; Jones et al., 1982, 1987). Our findings do not suggest that networks that process pitch and timing are identical, because some brain regions were more sensitive to pitch than to temporal repetition, and pitch and temporal repetition influenced behavioral performance differently; moreover, the two dimensions can be perceived separately by listeners (Thompson, 1994; Palmer & Krumhansl, 1987). Our findings suggest that similar motor networks are sensitive to repeated pitch and temporal structure when auditory sequences are transformed into motor sequences. Pitch and temporal structure may be processed more independently for different tasks (Bengtsson & Ullén, 2006). Peretz and Kolinsky (1993) suggested that pitch and temporal features are processed independently at early processing stages and integrated at later stages. Integrating pitch and temporal structures may be particularly advantageous when planning upcoming motor sequences is cognitively demanding, such as when performers must generate novel sequences (Berkowitz & Ansari, 2008) or when performers must plan entire movement sequences in advance, as in the current listen–playback task. In contrast, tasks that require less planning such as performing well-learned sequences (Bengtsson et al., 2004) or performing from musical notation (Bengtsson & Ullén, 2006) may entail more independent processing of pitch and temporal sequence structures. In the current task, both pitch and temporal structure were also present in each of the stimuli, which may have enhanced integration because pianists had to plan and execute movements based on both structures at once. Overall, the current findings suggest

that the motor system organizes responses based on multiple sensory cues and that this engages dPMC (Hoshi & Tanji, 2007), vPMC, pre-SMA, and parietal regions. Although the current study examined skilled performers, non-musicians may engage similar networks to produce pitch and temporal structure in auditory sequences; nonmusicians have recruited similar frontal–parietal networks as those described above during auditory–motor mapping tasks such as synchronizing with auditory rhythms (Chen et al., 2008b; Jäncke, Loose, Lutz, Specht, & Shah, 2000), listening to or silently performing musical sequences while imagining corresponding movements or sounds (Baumann et al., 2007), or learning to perform melodies by ear (Chen et al., 2012; Lahav et al., 2007). Therefore, our findings may generalize to nonskilled performers and potentially to other types of auditory–motor skills.

In summary, we have demonstrated that similar premotor and parietal networks are engaged in transforming pitch and temporal structures in music into motor movement, suggesting that the motor system processes pitch and temporal structure together. Parietal regions, IPS in particular, may specifically contribute to transforming pitch sequences into spatial coordinates for motor response. These findings contribute to our current knowledge of auditory–motor integration by demonstrating how motor regions respond to different levels of auditory sequence structure. The current findings suggest that much of the motor system is capable of processing multiple action-relevant stimulus features together, which may facilitate coordination of complex actions.

REFERENCES

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., et al. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *Neuroimage, 30,* 917–926.

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jäncke, L. (2007). A network for audio-motor coordination in

skilled pianists and non-musicians. *Brain Research, 1161,* 65–78.

Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., & Pike, B. (1999). Event-related fMRI of the auditory cortex. *Neuroimage, 10,* 417–429.

Bengtsson, S. L., Ehrsson, H. H., Forssberg, H., & Ullén, F. (2004). Dissociating brain regions controlling the temporal and ordinal structure of learned movement sequences. *European Journal of Neuroscience, 19,* 2591–2602.

Bengtsson, S. L., Ehrsson, H. H., Forssberg, H., & Ullén, F. (2005). Effector-independent voluntary timing: Behavioural and neuroimaging evidence. *European Journal of Neuroscience, 22,* 3255–3265.

Bengtsson, S. L., & Ullén, F. (2006). Dissociation between melodic and rhythmic processing during piano performance from musical scores. *Neuroimage, 30,* 272–284.

Berkowitz, A. L., & Ansari, D. (2008). Generation of novel motor sequences: The neural correlates of musical improvisation. *Neuroimage, 41,* 535–543.

Brown, S., & Martinez, M. J. (2007). Activation of premotor vocal areas during musical discrimination. *Brain and Cognition, 63,* 59–69.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008a). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18,* 2844–2854.

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

Chen, J. L., Rae, C., & Watkins, K. E. (2012). Learning to play a melody: An fMRI study examining the formation of auditory–motor associations. *Neuroimage, 59,* 1200–1208.

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

Cisek, P., & Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection action. *Neuron, 45,* 801–814.

Cusack, R. (2005). The intraparietal sulcus and perceptual organization. *Journal of Cognitive Neuroscience, 17,* 641–651.

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition, 74,* 1–32.

Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M.-H., Evans, A. C., Zilles, K., et al. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage, 36,* 511–521.

Foster, N. E. V., & Zatorre, R. J. (2010a). A role for the intraparietal sulcus in transforming musical pitch information. *Cerebral Cortex, 20,* 1350–1359.

Foster, N. E. V., & Zatorre, R. J. (2010b). Cortical structure predicts success in performing musical transformation judgments. *Neuroimage, 53,* 26–36.

Frey, S., Campbell, J. S. W., Pike, G. B., & Petrides, M. (2008). Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *The Journal of Neuroscience, 28,* 11435–11444.

Gaab, N., Gabrieli, J. D. E., & Glover, G. H. (2007). Assessing the influence of scanner background noise on auditory processing. I. An fMRI study comparing three experimental designs with varying degrees of scanner noise. *Human Brain Mapping, 28,* 703–720.

Garraux, G., McKinney, C., Wu, T., Kansaku, K., Nolte, G., & Hallett, M. (2005). Shared brain areas but not functional connections controlling movement timing and order. *The Journal of Neuroscience, 25,* 5290–5297.

Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage, 9,* 416–429.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15,* 20–25.

Grafton, S. T., Fagg, A. H., & Arbib, M. A. (1998). Dorsal premotor cortex and conditional movement selection: A PET functional mapping study. *Journal of Neurophysiology, 79,* 1092–1097.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience, 19,* 893–906.

Grefkes, C., Ritzl, A., Zilles, K., & Fink, G. R. (2004). Human medial intraparietal cortex subserves visuomotor coordinate transformation. *Neuroimage, 23,* 1494–1506.

Grefkes, C., Weiss, P. H., Zilles, K., & Fink, G. R. (2002). Crossmodal processing of object features in human anterior intraparietal cortex: An fMRI study implies equivalencies between humans and monkeys. *Neuron, 35,* 173–184.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10,* 14–23.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition, 92,* 67–99.

Hollinger, A. (2008). *Design of fMRI-compatible electronic musical interfaces* (Unpublished masters thesis). McGill University, Montreal.

Hollinger, A., Steele, C., Penhune, V., Zatorre, R., & Wanderley, M. (2007). fMRI-compatible electronic controllers. *Proceedings of the 2007 International Conference on New Interfaces for Musical Expression (NIME07), New York City, U.S.A.* (pp. 246–249).

Hoshi, E., & Tanji, J. (2007). Distinctions between dorsal and ventral premotor areas: Anatomical connectivity and functional properties. *Current Opinion in Neurobiology, 17,* 234–242.

Husain, M., & Nachev, P. (2007). Space and the parietal cortex. *Trends in Cognitive Sciences, 11,* 30–36.

Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience, 6,* 682–687.

Jäncke, L., Loose, R., Lutz, K., Specht, K., & Shah, N. J. (2000). Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Cognitive Brain Research, 10,* 51–66.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17,* 825–841.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis, 5,* 143–156.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics, 41,* 621–634.

Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics, 32,* 211–218.

Jones, M. R., Summerell, L., & Marshburn, E. (1987). Recognizing melodies: A dynamic interpretation. *The Quarterly Journal of Experimental Psychology, 39,* 89–121.

Jordan, K., Heinze, H.-J., Lutz, K., Kanowski, M., & Jäncke, L. (2001). Cortical activations during the mental rotation of different visual objects. *Neuroimage, 13,* 143–152.

Kostopoulos, P., & Petrides, M. (2003). The mid-ventrolateral prefrontal cortex: Insights into its role in memory retrieval. *European Journal of Neuroscience, 17,* 1489–1497.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network

while listening to newly acquired actions. *The Journal of Neuroscience, 27,* 308–314.

Lewis, P. A., Wing, A. M., Pope, P. A., Praamstra, P., & Miall, R. C. (2004). Brain activity correlates differentially with increasing temporal complexity of rhythms during initialisation, synchronisation, and continuation phases of paced finger tapping. *Neuropsychologia, 42,* 1301–1312.

Meister, I. G., Krings, T., Foltys, H., Boroojerdi, B., Müller, M., Töpper, R., et al. (2004). Playing piano in the mind—An fMRI study on music imagery and performance in pianists. *Cognitive Brain Research, 19,* 219–228.

Nachev, P., Wydell, H., O'Neill, K., Husain, M., & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage, 36,* T155–T163.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage, 25,* 653–660.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48,* 115–138.

Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 116–126.

Peretz, I., & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. *The Quarterly Journal of Experimental Psychology, 46,* 301–325.

Petrides, M., Alivisatos, B., & Evans, A. C. (1995). Functional activation of the human ventrolateral frontal cortex during mnemonic retrieval of verbal information. *Proceedings of the National Academy of Sciences, U.S.A., 92,* 5803–5807.

Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 949–964.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience, 12,* 718–724.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., et al. (1999). Neural representation of a rhythm depends on its interval ratio. *The Journal of Neuroscience, 19,* 10074–10081.

Sakai, K., Hikosaka, O., & Nakamura, K. (2004). Emergence of rhythm during motor learning. *Trends in Cognitive Sciences, 8,* 547–553.

Sakai, K., Ramnani, N., & Passingham, R. E. (2002). Learning of sequences of finger movements and timing: Frontal lobe and action-oriented representation. *Journal of Neurophysiology, 88,* 2035–2046.

Schubotz, R. I., & von Cramon, D. Y. (2001). Interval and ordinal properties of sequences are associated with distinct premotor areas. *Cerebral Cortex, 11,* 210–222.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17,* 143–155.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23,* S208–S219.

Thompson, W. F. (1994). Sensitivity to combinations of musical parameters: Pitch with duration, and pitch pattern with durational pattern. *Perception & Psychophysics, 56,* 363–374.

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for fMRI group analysis using Bayesian inference. *Neuroimage, 21,* 1732–1747.

Zacks, J. M. (2008). Neuroimaging studies of mental rotation: A meta-analysis and review. *Journal of Cognitive Neuroscience, 20,* 1–19.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience, 8,* 547–558.

Zatorre, R. J., Halpern, A. R., & Bouffard, M. (2010). Mental reversal of imagined melodies: A role for the posterior parietal cortex. *Journal of Cognitive Neuroscience, 22,* 775–789.

This article has been cited by:

1. Brian Mathias, William Gehring, Caroline Palmer. 2019. Electrical Brain Responses Reveal Sequential Constraints on Planning during Music Performance. *Brain Sciences* **9**:2, 25. [Crossref]

2. Rachel M. Brown, Virginia B. Penhune. 2018. Efficacy of Auditory versus Motor Learning for Skilled and Novice Performers. *Journal of Cognitive Neuroscience* **30**:11, 1657-1682. [Abstract] [Full Text] [PDF] [PDF Plus]

3. Lucía Vaquero, Neus Ramos-Escobar, Clément François, Virginia Penhune, Antoni Rodríguez-Fornells. 2018. White-matter structural connectivity predicts short-term melody and rhythm learning in non-musicians. *NeuroImage* **181**, 252-262. [Crossref]

4. Chen-Gia Tsai, Tai-Li Chou, Chia-Wei Li. 2018. Roles of posterior parietal and dorsal premotor cortices in relative pitch processing: Comparing musical intervals to lexical tones. *Neuropsychologia* **119**, 118-127. [Crossref]

5. Markus Christiner, Susanne Reiterer. 2018. Early Influence of Musical Abilities and Working Memory on Speech Imitation Abilities: Study with Pre-School Children. *Brain Sciences* **8**:9, 169. [Crossref]

6. Indiana Wollman, Virginia Penhune, Melanie Segado, Thibaut Carpentier, Robert J. Zatorre. 2018. Neural network retuning and neural predictors of learning success associated with cello training. *Proceedings of the National Academy of Sciences* **115**:26, E6056-E6064. [Crossref]

7. Eckart Altenmüller, Shinichi Furuya. 2017. Apollos Fluch und Segen: Musizieren als Neuroplastizitätsmotor. *e-Neuroforum* **23**:2. . [Crossref]

8. Eckart Altenmüller, Shinichi Furuya. 2017. Apollos Gift and Curse: Making Music as a model for Adaptive and Maladaptive Plasticity. *e-Neuroforum* **23**:2. . [Crossref]

9. Jing Lu, Hua Yang, Hui He, Seun Jeon, Changyue Hou, Alan C. Evans, Dezhong Yao. 2017. The Multiple-Demand System in the Novelty of Musical Improvisation: Evidence from an MRI Study on Composers. *Frontiers in Neuroscience* **11**. . [Crossref]

10. Masumi Wakita. 2016. Interaction between Perceived Action and Music Sequences in the Left Prefrontal Area. *Frontiers in Human Neuroscience* **10**. . [Crossref]

11. Hanna Poikonen, Petri Toiviainen, Mari Tervaniemi. 2016. Early auditory processing in musicians and dancers during a contemporary dance piece. *Scientific Reports* **6**:1. . [Crossref]

12. Robert Harris, Peter van Kranenburg, Bauke M. de Jong. 2016. Behavioral Quantification of Audiomotor Transformations in Improvising and Score-Dependent Musicians. *PLOS ONE* **11**:11, e0166033. [Crossref]

13. Carlotta Lega, Marianne A. Stephan, Robert J. Zatorre, Virginia Penhune. 2016. Testing the Role of Dorsal Premotor Cortex in Auditory-Motor Association Learning Using Transcranical Magnetic Stimulation (TMS). *PLOS ONE* **11**:9, e0163380. [Crossref]

14. Brian Mathias, Barbara Tillmann, Caroline Palmer. 2016. Sensory, Cognitive, and Sensorimotor Learning Effects in Recognition Memory for Music. *Journal of Cognitive Neuroscience* **28**:8, 1111-1126. [Abstract] [Full Text] [PDF] [PDF Plus]

15. Sibylle C. Herholz, Emily B.J. Coffey, Christo Pantev, Robert J. Zatorre. 2016. Dissociation of Neural Networks for Predisposition and for Training-Related Plasticity in Auditory-Motor Learning. *Cerebral Cortex* **26**:7, 3125-3134. [Crossref]

16. Salomi S. Asaridou, Atsuko Takashima, Dan Dediu, Peter Hagoort, James M. McQueen. 2016. Repetition Suppression in the Left Inferior Frontal Gyrus Predicts Tone Learning Performance. *Cerebral Cortex* **26**:6, 2728-2742. [Crossref]

17. Avrum D. Hollinger, Marcelo M. Wanderley. 2015. The Design, Implementation, and Testing of a Sensorized MRI-Compatible Cello. *IEEE Sensors Journal* **15**:11, 6125-6134. [Crossref]

18. Robert Harris, Bauke M. de Jong. 2015. Differential parietal and temporal contributions to music perception in improvising and score-dependent musicians, an fMRI study. *Brain Research* **1624**, 253-264. [Crossref]

19. M. E. Klein, R. J. Zatorre. 2015. Representations of Invariant Musical Categories Are Decodable by Pattern Analysis of Locally Distributed BOLD Responses in Superior Temporal and Intraparietal Sulci. *Cerebral Cortex* **25**:7, 1947-1957. [Crossref]

20. R.J. Zatorre, A.R. Halpern. Music 551-557. [Crossref]

21. Rachel M. Brown, Robert J. Zatorre, Virginia B. Penhune. Expert music performance: cognitive, neural, and developmental bases 57-86. [Crossref]

22. Petr Janata. Neural basis of music perception 187-205. [Crossref]

23. Gonçalo Padrão, Virginia Penhune, Ruth de Diego-Balaguer, Josep Marco-Pallares, Antoni Rodriguez-Fornells. 2014. ERP evidence of adaptive changes in error processing and attentional control during rhythm synchronization learning. *NeuroImage* **100**, 460-470. [Crossref]

24. Mathilde Groussard, Fausto Viader, Brigitte Landeau, Béatrice Desgranges, Francis Eustache, Hervé Platel. 2014. The effects of musical practice on structural plasticity: The dynamics of grey matter changes. *Brain and Cognition* **90**, 174-180. [Crossref]

25. Giacomo Novembre, Peter E. Keller. 2014. A conceptual review on action-perception coupling in the musiciansâ€™ brain: what is it good for?. *Frontiers in Human Neuroscience* **8**. . [Crossref]

26. Katja Kornysheva, Jörn Diedrichsen. 2014. Human premotor areas parse sequences into their spatial and temporal features. *eLife* **3**. . [Crossref]

27. Jennifer Anne Bailey, Robert J. Zatorre, Virginia B. Penhune. 2014. Early Musical Training Is Linked to Gray Matter Structure in the Ventral Premotor Cortex and Auditory–Motor Rhythm Synchronization Performance. *Journal of Cognitive Neuroscience* **26**:4, 755-767. [Abstract] [Full Text] [PDF] [PDF Plus]

28. Avrum D. Hollinger, Marcelo M. Wanderley. MRI-compatible optically-sensed cello 1-4. [Crossref]